PCT/AU 99/01079

AUSTRALIA

4

REC'D 03 FEB 2000

WIPO PCT

**Patent Office**
**Canberra**

I, LEANNE MYNOTT, TEAM LEADER EXAMINATION SUPPORT AND SALES hereby certify that annexed is a true copy of the Provisional specification in connection with Application No. PP 7469 for a patent by COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION filed on 03 December 1998.

WITNESS my hand this
Seventeenth day of January 2000

L. Mynott

LEANNE MYNOTT
<u>TEAM LEADER EXAMINATION</u>
<u>SUPPORT AND SALES</u>

COMMONWEALTH OF AUSTRALIA
AUSTRALIA
PATENT OFFICE

**PRIORITY DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

THIS PAGE BLANK (USPTO)

AUSTRALIA

Patents Act 1990

**PROVISIONAL SPECIFICATION**

**Applicant(s):** COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION

**Invention Title:** CONTROL OF FLOWERING

The invention is described in the following statement:

## CONTROL OF FLOWERING

This invention relates to the control of flowering and reproduction in plants, and in particular to agents and methods for inducing or suppressing flowering. The invention provides isolated nucleic acid molecules which are useful for inducing flowering, particularly initiating early flowering, for delaying or suppressing flowering, or for manipulating the flowering period.

## DETAILED DESCRIPTION OF THE INVENTION

The initiation of flowering in a plant occurs in response to internal signals, such as physiological age or levels of plant growth regulators, or may result from changes in environmental conditions such as day length or low temperature. It is well known that in a variety of plant species a crucial factor is day length, also known as photoperiod. In many plant species, including several ecotypes of the widely-used model plant species *Arabidopsis thaliana*, flowering is promoted by long day photoperiod, or by a period of low temperature (vernalization) (Napp-Zinn, 1985).

Control of flowering in both horticultural and crop plants represents a major problem in the agricultural industry, and is also a problem in forestry. Significant losses in yield of plants may result if non-uniform flowering of plants occurs; this applies both to field-grown and to glasshouse-grown plants. The problem is particularly acute for field-grown plants, which are frequently exposed to abnormal or unseasonal conditions which may result in induction of flowering at an inappropriate time. Efficient plant production requires the synchronization of flowering time between pollen donor and pollen receptor plants, and is particularly important to maximize market opportunities for glasshouse-grown plants.

Currently-available methods for regulation of flowering in plants are expensive and labour-intensive, and require the use of plant growth regulators, and/or controlled planting regimes and controlled-environment growth conditions. Consequently there is a need in the art for more efficient, cost-effective methods for controlling flowering time. These methods are applicable to a variety of commercially-significant plants species, including both horticultural plants, particularly those used in the cut-flower industry, and vegetable, cereal and other crop plants.

A class of genes known as MADS box genes comprises a distinctive conserved DNA binding domain, known as the MADS box, which in certain cases has been demonstrated to bind to $CC(A/T)_6GG$ DNA motifs. The MADS box genes encode a class of transcription factors, which was first identified in yeast and in mammals. Subsequently similar transcription factors were identified in a range of plants, including *Arabidopsis thaliana*, *Antirrhinum majus*, tomato, tobacco, petunia, corn, *Pinus* species and *Eucalyptus* species. In plants, the MADS box genes have a "K domain", which resembles the coiled-coil domains of keratin proteins, which are implicated in protein/protein interactions, an intervening (I) domain, and a carboxy terminal (C) domain. In plants the principal role of MADS box genes is in specifying inflorescence meristem identity, and floral organ identity and development. Certain MADS in florescence box genes have also been implicated as having roles in root and vegetative development.

We have now identified nucleic acid sequences comprising a MADS box in the model plant *Arabidopsis thaliana* which play a role in the control of flowering time. The effect on flowering depends on the degree of expression of the nucleic acid sequences.

SUMMARY OF THE INVENTION

In a first aspect, the invention provides an isolated nucleic acid molecule which is capable of altering the flowering time of a plant.

In one preferred embodiment, the invention provides an isolated nucleic acid molecule which is capable of delaying the flowering of a plant. Preferably expression of the nucleic acid molecule in the plant, in the sense orientation under the control of a promoter sequence, is capable of delaying the flowering of the plant.

In a second preferred embodiment, the isolated nucleic acid molecule of the invention is capable of accelerating the flowering of a plant. Preferably expression of the nucleic acid molecule in the plant in the anti-sense orientation under the control of a promoter sequence is capable of accelerating the flowering of the plant.

Preferably the nucleic acid molecule of the invention comprises a nucleotide sequence corresponding to a *FLOWERING LOCUS F* (*FLF*) gene. The nucleic acid molecule may be a genomic DNA, a cDNA, or a messenger RNA.

More preferably the nucleic acid molecule comprises the nucleotide sequence set out in SEQ ID NO. 1 or SEQ ID NO. 2, or a nucleotide molecule capable of hybridizing thereto under at least low stringency hybridization conditions.

In a second aspect, the invention provides a vector comprising a nucleic acid molecule according to the invention. The vector may be a virus, bacteriophage, plasmid, or bacterium. In a particularly preferred embodiment, the vector is a T-DNA vector present in a bacterium of the genus *Agrobacterium*, in particular *Agrobacterium tumefaciens*.

In a third aspect, the invention provides a plant cell comprising the nucleic acid of the invention.

In a fourth aspect, the invention provides a plant comprising the nucleic acid molecule of the invention.

In a fifth aspect, the invention provides a method of isolating a nucleic acid molecule capable of altering the flowering time of a target plant, comprising the step of using a nucleic acid molecule of the invention, or a functional portion thereof, as a hybridisation probe or polymerase chain reaction (PCR) primer, and detecting hybridisation. Suitable methods are very well known in the art.

Genes controlling developmental stages in plants, such as the gene associated with the nucleic acid of the invention, are highly conserved during evolution. Consequently the nucleic acid molecules and the methods of the invention are applicable to all plant species, whether the species is monocotyledonous or dicotyledonous. Thus the invention is generally applicable to flowering plants, including but not limited to ornamental, horticultural, agricultural and tree species. Methods for introducing exogenous DNA into plants of all these types, and for *in vitro* culture of plant tissue and regeneration of plant cells or tissues into whole plants, are known in the art. Methods for further generation and selection of commercially useful cultivars are also well known. Depending on the type of plant, it may be desirable to accelerate flowering *ie.* to induce early flowering, to synchronise flowering, to delay flowering or to suppress flowering.

For example it is desirable to suppress or delay flowering in many vegetable plants, in pasture grasses such as rye grass, or in sugar cane. Acceleration of flowering by induction of early flowering is desirable in a number of crop species, such as cotton, and in horticultural species.

We have surprisingly found that flowering can be delayed in proportion to the degree of expression of the nucleic acid molecule of the invention, and that early

flowering can be induced by reducing the expression of this nucleic acid molecule.

Thus in a sixth aspect the invention provides a method of delaying flowering in a plant, comprising the step of introducing a nucleic acid molecule of the invention into cells of the plant, optionally such that expression of the nucleic acid molecule is under the control of an inducible promoter, and over-expressing the nucleic acid molecule. Preferably the promoter is a tissue-specific promoter.

Preferably flowering is delayed for at least five days, preferably for at least twenty days, and more preferably for at least thirty days beyond the normal flowering period. Most preferably flowering is delayed for at least forty to fifty days.

According to a sixth aspect, the invention provides a method of inducing early flowering in a plant, comprising the step of reducing the degree of expression of a nucleic acid molecule of the invention in the plant. The reduction may be effected by any convenient means, including but not limited to transformation of the target plant with an anti-sense nucleic acid sequence, disruption of the nucleic acid sequence using a transposable element or transposon, or by a procedure such as vernalisation. The person skilled in the art will readily be able to select the most suitable procedure for the particular plant species in question. Optionally the method of the invention may be supplemented by other treatments, such as an exogenous gibberellin.

Preferably flowering is at least five days earlier than the normal flowering period, more preferably at least ten days, and most preferably at least fifteen days earlier than the normal flowering period.

In a number of embodiments of the invention, the nucleic acid molecule of the invention is operably linked to a promoter sequence capable of regulating the expression of the nucleic acid molecule; more preferably the promoter

sequence is adapted to regulate expression in a eukaryotic cell, most preferably a plant cell. The nucleic acid molecule of the invention may also be operably linked to a transcriptional terminator sequence.

Suitable promoter sequences are well known in the art, and include but are not limited to the CaMV 35S promoter, a NOS promoter, the octopine synthetase (OCS) promoter, a subclover stunt virus promoter and the *Arabidopisis thaliana ubiquitin* gene promoter. The person skilled in the art will readily be able to selected the most suitable promoter for a given purpose. In particular, for some purposes an inducible promoter may be desirable, and these are also well known in the art. Suitable transcriptional terminator sequences active in plant cells are also well known, and may be of bacterial, fungal, viral, animal or plant origin.

Suitable transcriptional terminators particularly suitable for use in the present invention include the nopaline synthase (NOS) gene transcriptional terminator of *Agrobacterium tumefaciens*, the transcriptional terminator of the Cauliflower mosaic virus (CaMV) 35S gene, the *zein* gene transcriptional terminator from *Zea mays*, and the Rubisco small subunit (SSU) gene transcriptional terminator sequences or subclover stunt virus (SCSV) gene sequence transcriptional terminators.

The nucleic acid molecule of the invention may be introduced into a plant cell or tissue by any suitable means. A variety of methods for introducing exogenous DNA into plant tissue (transformation) are known. These include, but are not limited to, direct DNA uptake into protoplasts (Krens *et al*, 1982; Paszkowski *et al*, 1984), polyethyleneglycol-mediated uptake to protoplasts (Armstrong *et al*, 1990), electrophoresis (Fromm *et al*, 1985), microinjection of DNA (Crossway *et al*, 1986), microparticle bombardment of tissue explants or cells (Christou *et al*, 1988; Sanford, 1993), or T-DNA-mediated transfer from *Agrobacterium* to the plant tissue.

Representative T-DNA vector systems are described in the following references: An *et al* (1985); Herrera-Estrella *et al* (1983a, b); Herrera-Estrella *et al* (1985). These transformation methods are applicable to plant tissue culture, or may be employed with whole plants (*in planta* transformation). Again a person skilled in the art will be able to select the most suitable method for any given plant.

Any plant tissue capable of subsequent clonal propagation, whether by organogenesis or embryogenesis, may be transformed with a vector of the present invention. The particular tissue chosen will vary, depending on the clonal propagation systems which are most suitable for the species being transformed. Suitable tissue targets include whole plant, leaf discs, pollen, embryos, cotyledons, hypocotyls, megagametophytes, callus tissue, existing meristematic tissue (*eg* apical meristem, axillary buds, and root meristems), and induced meristem tissue (*eg* cotyledon meristem and hypocotyl meristem).

The vector of the invention may additionally comprise a dominant selectable marker to facilitate cell selection and plant breeding. A variety of suitable markers is known in the art, including but not limited to the *NPTII* gene, genes encoding resistance to an antibiotic such as hygromycin or ampicillin or to a herbicide such as phospinothricin or glyphosate; a gene encoding a polypeptide which confers stress tolerance, such as superoxide dismutase; or a visually-detectable marker, such as green fluorescent protein or β-glucuronidase. The person skilled in the art will readily be able to select the most suitable marker for use in a specific case.

The invention is applicable to any dicotyledonous or monocotyledonous plant species, including but not limited to decorative flower, vegetable, fruit, cereal, tree, and other flowering species. Preferably the plant is selected from the group consisting of chrysanthemum, rose, gerbera, carnation, soya bean, cotton, oil seed rape,

wheat, maize, rice, *Canola* species, *Linola* species, *eucalyptus* species, pine and poplar. Forest species are to be understood to be within the scope of the invention.

For the purposes of this specification it will be clearly understood that the word "comprising" means "including but not limited to", and that the word "comprises" has a corresponding meaning.

The term "flowering time" as used herein means the time at which floral meristem tissue is first visually detectable in the plant, for example by light microscopy or using the naked eye. The measured flowering-time includes the time taken for the occurrence of the cellular processes in the differentiation of a floral meristem and subsequent cell divisions which enable such visual means to be used. The term "flowering time" also includes the time taken for the transition from a vegetative meristem to a floral meristem to occur, as measured visually, following the induction of flowering in the plant by the application thereto of a specific chemical, physical or environmental stimulus, such as a plant growth regulator, photoperiod or temperature regime, including the vernalisation of the plant. Alternatively flowering may be induced in response to an internal development signal in the plant. Those skilled in the art will be aware of the specific nature of such chemical, physical or environmental stimuli or internal developmental signals.

"Altering the flowering time" means that the time period in which floral meristem tissue is first visually detected in a plant is increased, decreased, or otherwise modified or regulated. Thus, flowering may be delayed, accelerated, inhibited, suppressed, or synchronized.

The term "meristem" refers to plant tissue in which cells are undergoing, or are capable of undergoing, rapid mitotic division followed by differentiation into cell types which are capable of forming a primordium which develops into an organ such as a leaf, root, stem, floral bud or other plant organ.

"Vegetative meristem" refers to a meristem in which the differentiation process produces a cell type which develops into a vegetative organ or non-reproductive organ, such as a leaf, petiole, bract, stem or root.

"Floral meristem" refers to a meristem in which the differentiation process produces a cell type which develops into an inflorescence meristem, a secondary inflorescence meristem, a floral organ or sexual reproductive organ, in which the meristem or organ, when developed, may comprise both reproductive and non-reproductive tissues, including, but not limited to, anthers, stamens, stigmas, ovules, carpels, petals and sepals. "Bolt" refers to an inflorescence stem of a rosette plant, and "bolting" is the development of such a stem.

The term "derived from" means that a particular integer or group of integers has originated from a particular organism or species as specified herein, but has not necessarily been obtained directly from that source.

A "homologue" of a nucleotide sequence refers to an isolated nucleic acid molecule which is substantially the same as the nucleic acid molecule of the present invention or its complementary nucleotide sequence, despite the occurrence within the sequence of one or more nucleotide substitutions, insertions, deletions, or rearrangements.

An "analogue" of a nucleotide sequence means an isolated nucleic acid molecule which is substantially the same as a nucleic acid molecule of the present invention or its complementary nucleic acid, despite the occurrence of any non-nucleotide constituents not normally present in the isolated nucleic acid molecule, for example carbohydrates, radiochemicals including radionucleotides, reporter molecules including, but not limited to digoxigenin, alkaline phosphatase or horseradish peroxidase.

A "derivative" of a nucleotide sequence means any isolated nucleic acid molecule which contains significant

sequence similarity to the molecule or a part thereof. The person skilled in the art will appreciate that the nucleotide sequence of the present invention may be subjected to mutagenesis to produce one or more single or multiple nucleotide substitutions, deletions and/or insertions. Nucleotide insertional derivatives of the nucleotide sequence of the present invention include 5' and 3' terminal fusions, as well as intra-sequence insertions of single or multiple nucleotides or nucleotide analogues. Insertional nucleotide sequence variants are those in which one or more nucleotides or nucleotide analogues are introduced into a predetermined site in the nucleotide sequence of the sequence, although random insertion is also possible; suitable screening of the resulting product is performed. Deletional variants are characterised by the removal of one or more nucleotides from the nucleotide sequence. Substitutional nucleotide variants are those in which at least one nucleotide in the sequence has been removed, and a different nucleotide analogue inserted in its place.

Reference in this specification to a "gene" is to be understood in its broadest context, and includes:

(i) a classical genomic sequence comprising transcriptional and/or translational regulatory sequences and/or a coding region and/or non-translated sequences (*ie*. introns and 5'- and 3'-untranslated sequences);

(ii) mRNA or cDNA corresponding to the coding regions (*ie*. exons), optionally additionally comprising 5'- or 3'-untranslated sequences of the gene; or

(iii) an amplified DNA fragment or other recombinant nucleic acid molecule produced *in vitro*, and comprising all or a part of the coding region and/or 5' or 3'-untranslated sequences of the gene.

The term "gene" is also used to describe synthetic or fusion molecules encoding all or part of a functional product. A functional product is one which comprises a sequence of nucleotides or is complementary to

a sequence of nucleotides which encodes a functional polypeptide, in particular the *FLF* polypeptide of the invention or a homologue, analogue or derivative thereof.

In some of the examples herein the *FLF* gene is referred to as gene B. These two terms are synonymous.

## BRIEF DESCRIPTION OF THE FIGURES

Figure 1 is a photograph showing wild-type C24 (left), the late flowering T-DNA tagged *flf* mutant (middle) at 70 days after germination, and *flf* mutant at 150 days, showing the domed shape caused by vegetative bolts (right). The bar represents 5 cm.

Figure 2 shows segregation of two T-DNA inserts with the late flowering phenotype

**A.** Genomic DNA isolated from a $T_2$ population segregating for early (E), late (L) and very late (VL) flowering, digested with *Eco*RI and probed with the *NPT*II gene.

**B.** Physical map of the two T-DNA inserts linked to the *FLF* locus, showing their orientation. EcoRI sites are labelled RI; LB and RB represent the left and right borders, respectively, of the T-DNA. The triangle symbol represents the site of deletion of 30 bp to the right of the T-DNA. The arrows represent the direction of transcription of the genes.

**C.** Representation of a 27 kb region of *Arabidopsis* mutant DNA containing the gene A and gene B loci, showing the location of the T-DNA inserts. DNA fragments from the flanking plant DNA (probes 2 and 3) were used as probes to isolate cDNA clones. *Hind*III (H), *Bam*HI (B) and *Eco*RI (R1) restriction enzyme sites are indicated. The positions of gene A and gene B are shown, and their directions of transcription are indicated by arrows.

**D.** The 6.5kb and 6.8 kb *Bam*HI fragments, isolated from a genomic library of wild-type C24 with probes 1 and 2 from Figure 2C, spanning the site of the T-DNA insertions. Restriction sites are as in Figure 2C.

Figure 3 shows levels of expression of gene A (**A**) and gene B (**B**) in 30 day old wild-type C24 plants (lane 1), hemizygous (lane 2) and homozygous (lane 3) *flf* mutant plants.

Figure 4 is a photograph showing 35S::*FLF* T1 transgenic plants in C24 (left) and Landsberg *erecta* (right) ecotypes. The C24 transgenics were either early-flowering (back) or late-flowering (front). The Landsberg *erecta* transgenics were either late-flowering (left) or flowered at normal time (right).

Figure 5 shows the *FLF* gene structure and the expression pattern of the *FLF* transcript

**A**. Genomic structure of the *FLF* gene, showing location and site of introns and location of the MADS box, intervening domain (I), K domain (K) and carboxy terminal domain (C). The numbers below the line represent the number of base pairs in each exon.

**B**. Pattern of expression of *FLF* mRNA in C24 plants: roots (lane 1) and rosette leaves (veg rosette leaves, lane 2) from *in vitro* grown vegetative plants, cauline leaves (lane 3), bolt stems (stem, lane 4), floral apex and buds (buds, lane 5) from soil grown plants with bolt stems between 1 and 5 cm. Mature flowers (flowers, lane 6) and siliques (siliques, lane 7) were collected from older plants. Plants were grown under 16 h photoperiod conditions. The RNA gel blots for B-E were probed with a riboprobe transcribed from the *FLF* (Gene B) cDNA clone linearised to remove the MADS box region.

**C**. Expression level of *FLF* mRNA in whole C24 or *flf* mutant plants, harvested every 10 days (as indicated by the numerals) until the majority of the C24 plants were bolting (50 days under these growth conditions).

**D**. Expression of *FLF* mRNA in C24 (lane 1-6) and *flf* (lane 7-12) plants grown for 21 days in 8 h fluorescent photoperiod, and then at the end of the 21$^{st}$ photoperiod either kept in the same conditions (SD; lane 1, 2, 7, 8) or transferred to continuous dark (CD; lanes 3, 4,

9, 10) or continuous light (CL; lanes 5, 6, 11, 12). Plants were harvested either just prior to what would have been the start of the following photoperiod (dawn; lane 1, 3, 5, 7, 9, 11), or just prior to the end of the photoperiod (dusk; lanes 2, 4, 6, 8, 10, 12). Transcript levels were a little higher at the start of the photoperiod, but this pattern was not altered in the mutant.

**E.** Effect of gibberellic acid ($GA_3$) treatment and vernalization on the *FLF* transcript in C24 (lanes 1-3) and *flf* mutant (lanes 4-6) seedlings. RNA was isolated from 12 day old seedlings that had either had no treatment (C; lanes 1 and 4), been grown on medium containing $10^{-5}$ M $GA_3$ (G; lanes 2 and 5) or had a pretreatment of 3 weeks at 4°C (V; lanes 3 and 6).

Figure 6 shows:

**A.** Genomic DNA isolated from individual *flf*, efSL3 (M2), efSL4 (M2) and C24 plants digested with EcoRI and probed with a probe directed to the 3' region of Ac. The DNA for the *flf* sample was extracted from plants which contained a third T-DNA band, hence the band at about 8 kb. The presence of this third band had no effect on flowering time, and is therefore irrelevant.

**B.** As for **A**, except that the probe was probe 4 (see Figure 2C). The DNA for the efSL3 (M2) and efSL4 (M2) samples was extracted from bulked M2 plants which contained neo-lates as well as early-flowering mutants. Therefore there is some of the 2.7 kb band present in these DNA extracts. Other DNA isolated from individual early-flowering plants does not contain a band at 2.7 kb.

**C.** Location of Ac insertion in intron I

**D.** Expression level of *FLF* gene in 15 day old rosette leaves of C24 (lane 1), *flf* (lane2), ef SL3 (lane 3), ef SL4 (lane 4). The M2 early flowering mutants had just started to bolt, whereas the other plants remained vegetative.

Figure 7 shows gel blots from neo-late plants.

**A.** Genomic DNA was isolated from the 6 neo-late mutant plants and from efSL3 (M2), efSL4 (M2), flf and C24 plants, and digested with EcoRI. The DNA gel blot was probed with the 3' region of Ac.

**B.** Total RNA was isolated from a mixture of rosette and cauline leaves from the 6 neo-late plants, *flf* and C24. The RNA gel blot was probed as in Figure 5.

Figure 8 shows expression of the FLF gene in ecotypes and late-flowering mutants.

Total RNA was extracted from 12 d old seedlings and RNA gel bolts were probed as in Figure 5.

**A.** Expression in a number of different Arabidopsis ecotypes

**B.** Expression in Landsberg *erecta* (L.*er.*)and Landsberg *erecta* lines which contain late alleles at either the *FRI* (L.*er.*-$FRI^{Sf2}$) or FLC (L.*er.*-$FLC^{Sf2}$, L.*er.*-$FLC^{Col}$) loci.

**C.** Expression in late-flowering mutants in either L.*er.* ecotype (*fca, fve, fpa, gi, co, fha, fwa, fd, fe, ft*) or Ws (*ld*). The mutants *vrn1* and *vrn2* were isolated in the *fca* background, and only *vrn1* has been segregated away from the *fca* mutant locus.

## DETAILED DESCRIPTION OF THE INVENTION

The invention will now be described in detail by way of reference only to the following non-limiting examples, and to the figures.

## GENERAL METHODS

### Plant Material and Growth Conditions

*Arabidopsis* was grown either in pots containing a mixture of 50% sand and 50% loam, or aseptically in test-tubes or petri dishes containing a modified Murashige and Skoog (MS) medium (Langridge, 1957). Unless otherwise stated, all plants were grown in artificially lit cabinets at 21° or 23°C, under long day (16 hr light, 8 hr dark)

conditions using cool white fluorescent lights at an intensity of 200 μM $m^{-2}s^{-1}$.

Plants were vernalized by germinating seed in the dark for either 3, 4 or 8 weeks at 4°C. Following this cold treatment, seedlings were transferred to long day photoperiods at 23°C and times to flower, measured as the time until stem elongation (bolting) was observed, were determined, beginning from the first day at the higher temperature.

***Arabidopsis* Transformation**

*Arabidopsis* was transformed either by root transformation (Valvelkens *et al.*, 1988) for the generation of the *flf* mutant and gene A transgenic plants, or by *in planta* transformation (Bechtold *et al*, 1993) for the gene B transgenic plants. The late-flowering mutant (*flf*) arose during transformation of the early-flowering ecotype C24 with a modified binary vector pBinΔAc (Finnegan *et al.*, 1993). This vector contains the neomycin phosphotransferase II (*NPTII*) gene under the control of the nopaline synthase (*NOS*) promoter, together with a deleted maize Ac transposable element inserted within the untranslated leader of the β-glucuronidase gene, in the reverse orientation to the direction of transcription.

Example 1 Identification of a Mutant Locus Associated With a Late Flowering Phenotype

Following transformation of the *Arabidopsis thaliana* ecotype C24 with a T-DNA construct containing an Ac transposable element, individual $T_0$ plants, resistant to kanamycin, were allowed to self-pollinate, and the $T_1$ progeny screened for families that flowered significantly later than parental C24 plants.

Some of the plants of family 14-58 flowered after 70 days compared to 30 days for the C24 control plants. Segregation analysis of the progeny of a selfed late-flowering $T_1$ plant from family 14-58, showed 53 "late"

(flowering time >70 days) compared to 15 "early" flowering plants (flowering time 30 days). The result fits a 3:1 segregation ratio ($\chi^2$=0.313 P>0.5), and is consistent with the late-flowering phenotype being a consequence of a single mutation. We have designated the mutant locus *FLOWERING LOCUS F* (*FLF*).

Within the segregating progeny, the late-flowering plants could be further differentiated into two classes; "lates", flowering between 70 to 90 days and "very lates", which flowered later than 150 days. Some of the "very late" plants had not flowered after a year of growth. The "very late" *flf* mutant, shown in Figure 1, produced leaves at a rate similar to non-transformed C24, and had many more leaves at flowering than C24. After several months of growth, bolts arose from the internodes between rosette leaves. These bolts elongated approximately two to three cm and formed aerial rosette structures, giving the mutant plants a dome-like appearance, similar to that described for the *fld* mutant (Chou and Yang, 1998).

The late-flowering phenotype observed is more extreme than any of the previously-reported late-flowering mutants and ecotypes (Koornneef *et al*, 1991). Progeny tests showed that selfed "very lates" produced only "very late" progeny, whereas selfed "late" plants segregated 1:2:1 for very late, late and early flowering plants. This segregation pattern is consistent with a semi-dominant mutation, with the lateness in flowering being proportional to the level of gene product.

## Example 2 Construction and Screening of Genomic Libraries

A genomic library of the *flf* mutant was constructed by partial digestion of total plant DNA with the restriction enzyme *Sau*3AI and ligation into the phage vector λEMBL4. The resulting library was screened using a $^{32}$P-dCTP labelled probe of the *NPT*II gene (Feinberg and Vogelstein, 1983). Four positive phage clones were

purified and restriction mapped. Together these spanned 27kb of plant DNA flanking the site of insertion of T-DNA. A 2.3kb *Bam*HI-*Eco*RI and a 2.7 kb *Eco*RI fragment (probes 1 and 2 respectively, Figure 2C) isolated from this flanking plant DNA were subsequently used to probe a genomic library of wild-type *Arabidopsis* C24, made from *Bam*HI-digested DNA and cloned into λEMBL4. Probe 1 hybridized to a genomic clone containing 6.5kb of plant DNA spanning the T-DNA insertion site, and probe 2 hybridized to a genomic clone containing 6.8 kb of adjacent sequence.

Example 3 Isolation of the *FLF* Gene

*Two T-DNAs segregate with the very late flowering phenotype*

The very late-flowering phenotype segregates with two T-DNA bands identifiable by Southern analysis, which were designated bands 1 and 5. Southern blotting showed that bands 1 and 5 are inverted and adjacent:

RB<—— LB LB——>RB

1 5

(LB, left border; RB, right border).

The size of the bands, combined with sequence analysis, places the smaller band (band 5) closest to the *FLF* gene. Recombinant inbred lines were used to map the *FLF* region to the top of chromosome 5, 4 cM from RFLP marker 447. This places the *FLF* gene near *FLC*, a gene known to control flowering time in ecotypes of *Arabidopsis*.

A 27 kb segment of genomic DNA from the mutant around the site of T-DNA insertion was mapped. C24 genomic clones covering 13 kb around the T-DNA insertion site in the mutant were sequenced. Two probes, on either side of the T-DNA insert, were used to screen cDNA libraries. A 4.6 kb EcoRI/BamHI fragment from a C24 genomic clone (containing 4.4kb of sequence "upstream" of the T-DNA insertion site and 0.2 kb "downstream" of the T-DNA insertion site) was used to isolate cDNA clones identifying

"gene A" as a transcribed region. A 2.7 kb EcoRI fragment, covering the region 0.4 kb "downstream" to 3.1 kb downstream of the T-DNA inserts, was used to isolate cDNA clones identifying "gene B" as a transcribed region.

Comparison of part of the intergenic region sequence between ecotypes C24 and Ws revealed an insertion into the C24 DNA of an approximately 200 bp sequence, 420 bp to the "right" of the stop codon of gene A and 120 bp to the "left" of the T-DNA insertion. The sequence is present in both C24 wild-type and the *flf* mutant. The 200 bp insertion shows 100 % homology to ORF167 of the *Arabidopsis* mitochondrial genome. By PCR analysis we have determined that this sequence is absent in Landsberg *erecta* and Columbia ecotypes. The significance of this inserted DNA segment is unknown.

A map of the overall region is shown in Figure 2B.

<u>Example 4</u>　　<u>cDNA Libraries</u>

In order to identify expressed genes closely linked to the T-DNA, three *Arabidopsis* cDNA libraries (Elledge *et al*, 1991; Weigel *et al*, 1992; Newman *et al*, 1994) were screened, using probes to plant DNA around the T-DNA insertion site (probes 2 and 3, Figure 2C). Two classes of cDNA clones were isolated. These were respectively designated gene A and gene B. Gene B was subsequently re-designated as *FLF*. Two gene A cDNAs were isolated with a 4.6 kb EcoRI-BamHI fragment (probe 3, Figure 2C) from a screen of 200,000 λYes clones; however, no clones were isolated with the 2.7 Kb EcoRI fragment (probe 2, Figure 2C). The gene A cDNAs were subcloned from the phage by site-specific recombination, using the CRE protein (provided by the *E. coli* strain BNN132) and the lox sites within the vector (Elledge *et al*, 1991). The larger, almost full-length gene A cDNA was further subcloned into pBluescript SK(-) (Stratagene). Full-length gene A cDNAs were subsequently isolated by screening a Landsberg *erecta*

flower cDNA library (Weigel *et al*, 1992). The mutant and wild-type genomic clones corresponding to the isolated cDNA were also subcloned as smaller fragments into pBluescript SK(-). As no cDNA clones were isolated with probe 2 (Figure 2C) from either cDNA library, a third library was screened. Four full-length gene B cDNA clones were isolated from a λPRL2 cDNA library derived from different tissues and developmental stages (Newman *et al*, 1994). All cDNAs and the mutant and wild-type genomic clones were sequenced on both strands by the dideoxy chain termination method (Sanger *et al*, 1977) using fluorescent primers (Brumbaugh *et al*, 1988). The University of Wisconsin GCG software package was employed for sequence analysis (Devereux *et al*, 1984). The nucleotide and predicted protein sequences were used to search the GenBank database for any homologous sequences.

Example 5 Construction of 35S::gene A plasmid

As the larger of the initially isolated gene A cDNA clones lacked the AT of the ATG of the start codon, oligonucleotide-directed mutagenesis was employed to generate a 200bp fragment from the 5′ end of the cDNA which contained the absent nucleotides. Two oligonucleotides were synthesized on an Applied Biosystems DNA Synthesizer for this purpose:-

(1) 5' AAG**CCGCGG**ACAATGGAAGCTGTAAGATGC 3'

(2) 5' GAGAGGCTG**GTTAAC**CGGAG 3'.

The nucleotides indicated in bold show the locations of *Sac*II and *Hpa*I restriction sites within the primers. The amplification reaction was carried out in a 10 µl final volume that contained 2 µM of each oligonucleotide primer, 200pg of *Hind*III-cleaved cDNA as a template, 0.2 units of Taq polymerase and 125 µM of each of the four deoxynucleotides. Conditions for the amplification were as follows: 95°C for 2 mins, 5 cycles

consisting of 15 s denaturation 95°C, annealing at 40°C for 30 s, and polymerization at 72°C for 1 min, followed by 25 cycles where the annealing temperature was raised to 50°C for 15 sec and finally 30°C for 1 min. The resulting 200 bp PCR fragment was cloned into *Sac*II and *Hpa*I sites in the original cDNA plasmid, and then sequenced to ensure that no mutations had been introduced during the amplification procedure. Sense binary constructs were made by digesting the full length cDNA with *Eco*RI and *Sac*II, end filling the recessed termini using the Klenow fragment of DNA polymerase I, and ligating the released 1.4 kb fragment into the *Sma*I site of the expression vector pDH51 (Pietrzak *et al*, 1986). This places the expression of the *FLF* cDNA under the control of the CaMV 35S promoter. Recombinant plasmids, containing the cDNA in the desired orientation, were cleaved with *Eco*RI and cloned between the right and left border sequences of the binary vector pBin19 (Bevan, 1984). The binary construct was transferred to *Agrobacterium tumefaciens* strain AGL1 (Lazo *et al*, 1991) by triparental mating, employing pRK2013 as the helper plasmid. Roots of wild-type C24 plants were transformed (Valvelkens *et al*, 1988) using the *NPT*II gene as a selectable marker to identify transgenic plants.

Example 6 Construction of 35S::gene B plasmid

A binary construct containing gene B under the control of the CaMV 35S promoter was generated by cloning a XhoI/SpeI digested PCR product, amplified using the gene B cDNA clone as template with the primers, using methods similar to those described in Example 5:

5' CCG**CTCGAG**<u>CTTAGTATCTCCGGCG</u> 3' and
5' GG**ACTAGT**<u>CGCCCTTATCAGCGGA</u> 3',

in which restriction sites are shown in bold, and the sequence hybridizing to gene B cDNA is underlined, into XhoI/SpeI digested pART7 (Gleaves, 1992) containing the

CaMV 35S promoter. The 35S::gene B cassette was then subcloned using NotI into pART27 (Gleaves, 1992) and introduced into *A. tumefaciens* strain GV3101 (Koncz and Schell, 1986) as described above. Transgenic plants were generated by *in planta* transformation (Bechthold et al, 1993)

Example 7 DNA Gel Blot Analysis

Total genomic DNA was isolated either by following the cetyltrimethylammonium bromide (CTAB) procedure (Dean *et al*, 1992), or as described by McNellis et al. (1998). 2-3 μg of DNA was digested with the appropriate restriction enzyme, electrophoresed on an 0.8% agarose gel, and blotted onto Hybond N+ membranes (Southern, 1975). Probes were labelled with $^{32}$P-dCTP using the random primer method (Feinberg and Vogelstein, 1983).

The NPTII probe was generated as described above. The 3'Ac probe was a SphI fragment (Lawrence *et al*, 1993), and probe 4 was generated by amplification of the wild-type genomic clone with primers:

5'-GTATAGGGCACATGCCC-3' and
5'-CACTCGGAGCTGTGCC-3'.

This results in a 570 bp subset of probe 2 sequence, lacking the MADS box, to eliminate cross-hybridization.

Example 8 RNA Extraction and RNA Gel Blot Analysis

Total RNA was extracted from approximately 1g of plant tissue, following the method of Longemann *et al.* (1987). 10-20 μg of total RNA was electrophoresed on 2.2 M formaldehyde/agarose gels, and blotted onto Hybond N nylon filters. T7 or SP6 polymerase transcription of the linearised gene A or gene B plasmid template (containing the complete cDNA for gene A, linearised to remove MADS box for gene B) was used to generate antisense $^{32}$P-dUTP

labelled riboprobes. Filters were hybridized as described by Dolferus *et al* (1994), washed with a final solution of 0.1xSSC, 0.1% SDS at 65°C. For gene A Northern blots it was necessary to treat filters with RNase A as previously described (Dolferus *et al,* 1994) to avoid ribosomal trapping. The filters were exposed to phosphor screens for quantification of signal intensity using a phosphorimager (Molecular Dynamics, USA). RNA size markers were used to determine the size of the gene A and gene B transcripts.

Example 9 RFLP Mapping

Sixty-four recombinant inbred lines (Lister and Dean, 1993) were digested with *Bam*HI, and Southern blots were probed with gene A. The Mapmaker program was employed to compare the data with RFLP data for 68 mapped markers. Fine mapping of gene A was performed using DNA from F2 plants generated from a cross between Landsberg *erecta* and the *flf* mutant. A *Hpa*I digest of 62 $F_2$ DNA was probed with the chromosome 5 RFLP marker 447 (Chang *et al.*, 1988). A restriction fragment length polymorphism (RFLP) was found between the parental lines, Landsberg *erecta* and Columbia, by digestion of genomic DNA with *Bam*HI and probing with the gene A cDNA.

Example 10 Insertion Of Two Inverted and Adjacent T-DNAs Produces A Partial-Dominant Late Flowering Mutant

DNA gel blot analysis of a late flowering $T_1$ parent showed five T-DNA inserts. Figure 2A shows a DNA gel blot of 16 segregating progeny plants of a selfed, hemizygous mutant plant derived from this $T_1$ parent. In a total of 70 progeny, only two inserts segregated with the late-flowering phenotype (bands 1 and 5); "late" plants were hemizygous for these two T-DNAs, and "very late" plants were homozygous for the same two inserts. The two linked T-DNAs were segregated away from the other T-DNA

inserts by backcrossing to non-transformed C24. Plants containing only the two T-DNA inserts (Figure 2A, bands 1 and 5) were identified by DNA gel blot analysis, and further analysis showed that the two inserts were adjacent and in inverted orientation (Figure 2B). DNA gel blot analysis, using probes derived from segments of the T-DNA construct and Ac demonstrated that no movement of the Ac transposable elements had occurred.

A genomic DNA library from a late flowering plant containing only the two linked T-DNA inserts was screened with an *NPT*II probe to isolate DNA segments spanning the site of T-DNA insertion. Three overlapping clones were isolated from the left side of the T-DNA. The longest of these, together with one clone isolated from the right side, is depicted in Figure 2C. These clones accounted for a total of 27 kb of plant DNA spanning the site of the T-DNA insertion. C24 genomic clones were isolated from a genomic DNA library, which was prepared by digestion of total *Arabidopsis* DNA with the restriction enzyme *Bam*HI and ligation of the digest into the phage vector λEMBL4, using probes 1 and 2 (Figure 2C). Clones containing a 6.5kb *Bam*HI fragment spanning the insertion site and a 6.8 kb BamHI fragment downstream of the insertion site were characterized (Figure 2D). The genomic DNA sequence, cDNA sequence, and predicted protein sequence are set out in SEQ ID NO. 1, SEQ ID NO. 2 and SEQ ID NO. 3 respectively. The genomic DNA sequence includes about 2 kb of promoter sequence, and 6 introns.

The chromosomal location of the *FLF* region was determined using 64 F9 recombinant inbred lines and data for several known markers (Lister and Dean, 1993). The Mapmaker program located the *FLF* region to the top of chromosome 5, 4cM from the RFLP marker 447, placing the gene in the vicinity of three other genes, *TFL*, *FLC* and *EMF*1, which are involved in the regulation of flower initiation (Shannon and Meeks-Wagner, 1991; Lee *et al*, 1994b; Koornneef *et al*, 1994; Sung *et al*, 1992).

Example 11 Two Genes Flanking the T-DNA Inserts Have Increased Expression In The *flf* Mutant

Transcriptionally active regions in the plant DNA spanning the site of T-DNA insertion were identified by screening a cDNA library with probes derived from either side of the T-DNAs (probes 2 and 3, Figure 2C). Two overlapping partial-length cDNA clones were isolated with probe 3; full-length clones containing inserts of approximately 1.5 kb were isolated by screening a second cDNA library. Four identical full-length cDNA clones containing inserts of 1.0 kb were isolated with probe 2. Comparison of cDNA and genomic sequence revealed that the T-DNAs had inserted between the two transcribed regions, 591 bp downstream of the polyadenylation site of gene A and 2.3 kb upstream of the start codon of gene B. Comparison of the mutant and C24 genomic sequences revealed a 30 bp deletion immediately downstream of the insertion site, with no further differences identifiable.

As neither gene was disrupted by the insertion of the T-DNAs, we investigated whether the expression of the genes was altered in the mutant. Figure 3A shows an RNA gel blot performed using RNA isolated from 30 d old leaf tissue of ecotype C24, and hemi- and homozygous *flf* mutant plants grown under identical conditions. As shown in Figure 3 antisense riboprobes specific for either gene A or gene B revealed that both genes are more highly expressed in the *flf* homozygote leaf tissue than in ecotype C24, with the 1.5 kb gene A transcript being approximately 10 times more highly expressed in the mutant and the 1.0 kb gene B transcript being approximately two-fold overexpressed in the homozygous mutant. The hemizygous mutant has an intermediate level of expression of both genes. Without wishing to be limited by any proposed mechanism, we believe that insertion of the T-DNA complete with Ac elements has caused this over-expression.

Example 12 Transgenic Plants Over-Expressing Gene B Have Altered Flowering Time

In order to determine which gene is responsible for the late-flowering phenotype, we transformed C24 with constructs containing either gene under the control of the CaMV 35S promoter. 49 transgenic lines were generated with the gene A construct, and flowering time was assessed in the T2 generation. The majority of the transgenic lines showed no variation from wild-type flowering time; however, a few lines were slightly late-flowering. As shown in Table 1, four of the transgenic lines were significantly later flowering than the C24 wild-type. However, there was no correlation between time to flowering and the level of gene A expression. 23 transgenic lines were generated with the gene B construct, and differences in flowering time compared to C24 were apparent in the T1 generation, *ie*. in the primary transformants. 17 T1 plants flowered earlier (range 15-25 d) than non-transformed C24 (30 d) under these conditions; eight of these showed either full or partial sterility. Four flowered at around the same time as C24, and two had not flowered after 90 d. Examination of the level of gene B mRNA transcript in kanamycin-resistant progeny of two early-flowering T1 plants, and rosette leaves from two late-flowering T1 plants, revealed a high level of gene B expression in all the transgenic plants.

Table 1

Flowering time of 35S::gene A transgenic plants

| | Flowering time | Relative expression level of gene A mRNA |
|---|---|---|
| C24 | 29.9 ± 0.6 | 1 |
| B2 | 38.3 ± 0.9 | 2 |
| A53 | 35.8 ± 0.8 | 20 |
| A54 | 38.4 ± 0.5 | 8 |
| A93 | 35.4 ± 0.5 | 4 |
| *flf* | >>50 | 10 |

Transgenic and *flf* mutant seeds were germinated on MS plates + 50 µg/ml kanamycin, and C24 seeds were germinated on MS plates. At least 20 12 d seedlings were transplanted into individual soil pots and grown at 23°C under fluorescent lights (16 h light, 8 h dark). Flowering time was recorded as the number of days to stem elongation. Total RNA was extracted from 14 d *in vitro* grown seedlings and used for quantitation of *FLF* expression levels.

Flowering time of 35S::gene B T1 transgenic plants

| | Flowering time | Relative expression level of gene B mRNA |
|---|---|---|
| C24 | 30 d | 1 |
| B4 | 18 d | > 10 |
| B5 | 18 d | > 10 |
| B11 | > 80 d | > 10 |
| B12 | > 80 d | > 10 |
| Landsberg *erecta* | 20 d | n.d. |
| B36 | > 80 d | > 10 |
| B45 | > 80 d | > 10 |
| *flf* | > 80 d | 2 |

Transformant seeds, harvested from *in planta* transformed plants, were selected on MS plates containing 50 µg/ml kanamycin. Kanamycin resistant T1 seedlings were transferred to soil at 20 d. The level of gene B transcript was determined in kanamycin resistant T2 plants for B2 and B5, and from young leaves from the T1 late-flowering plants (B11, B12, B36, B45). n.d.: not detectable.

It is surprising that over-expression of gene B gave two completely opposite phenotypes. In order to clarify this, we generated transgenics containing gene B under the control of the 35S promoter in ecotype Landsberg *erecta*. Of the 24 T1 lines generated, none flowered earlier than wild-type Landsberg *erecta*, 12 had not bolted after 70 days, and 3 bolted after about 40 days, compared with 25 days for Landsberg *erecta*. Two of the three lines which bolted after 40 days exhibited floral abnormalities and partial sterility. Total RNA was isolated from rosette leaves of two non-flowering T1 plants (B36, B45; Table 1), both of which had high expression levels of the transgene.

Therefore in Landsberg *erecta* over-expression of gene B causes a delay in flowering time, whereas in C24 it causes either a delay in flowering, or causes the plants to flower significantly earlier. This may be mediated by a dominant negative effect or by a form of post-transcriptional gene silencing. Analysis of protein expression levels is being pursued in order to clarify this point. Presumably a difference in the genetic background of the two ecotypes is responsible for the difference observed between ecotypes.

These results demonstrate that over-expression of gene B causes late-flowering, whereas gene A has little effect on flowering time, indicating that over-expression of gene B is the most likely cause of the late-flowering *flf* phenotype and that this gene encodes a dosage-dependent

repressor of flowering. Gene B will hereafter be referred to as *FLF*.

Example 13 Anti-Sense Constructs

Anti-sense plant constructs have been generated using an anti-sense *FLF* gene construct under the control of the CaMV 35S promoter. A 35S::*FLF* antisense binary construct was generated by cloning the EcoRI/SpeI digested PCR product amplified with primers

CG**GAATTC**TCACACGAATAAGGTAC and

GG**ACTAGT**GGTCAAGATCCTTGATC

as described for the 35S::*FLF* construct. This amplified the region downstream of the MADS box, so that the antisense construct lacks the MADS box region. The PCR product was cloned into pART7 and pBART 27 (which is a derivative of pART27), and transgenic plants were generated as described above, except that *Bar* gene was used as the selectable marker.

25 T1 C24 transgenic plants were generated with a construct in which the 3' end of the *FLF* gene, in the antisense orientation, was under the control of a 35S promoter (35S::*FLF*AS). Approximately half of the T1 plants had flowered before 20 days of growth, compared to 30 days for the non-transformed strain

Transgenic plants were produced in the C24 and the Columbia ecotypes. Of the six T1 plants produced in the Columbia ecotype, three bolted earlier than wild-type Columbia. Wild-type C24 plants bolt at about 30 days, and wild-type Columbia plants bolt at about 20 days.

These results indicate that antisense acts to decrease flowering time, presumably by decreasing the level of the *FLF* transcript.

Example 14 The *FLF* Gene Is A Novel MADS Box Gene

The *FLF* cDNA sequence has strong homology to a class of transcription factors known as MADS box genes. The *FLF* sequence shows greatest similarity to the MADS gene

AGL14 in the M-I-K domain, but over the entire cDNA sequence it shows greater similarity to *CAL* (*CAULIFLOWER*) and *AP*1 (*APETALA*1). The location of the MADS box, I domain, K domain and C terminal domain are indicated in Figure 5A. The K domain is typical of those of other plant MADS genes. Comparison of the genomic sequence (SEQ ID NO. 1) and cDNA sequence (SEQ ID NO. 2) of the *FLF* gene revealed the presence of 6 introns, with intron I being 3.5 kb. The predicted protein is 196 amino acids long (SEQ ID NO. 3), which is shorter than the proteins encoded by most MADS box genes.

One of the main roles of MADS box genes in plant development which has been described to date is in specifying floral organ identity. Other roles for MADS-box genes include specifying root architecture and vegetative growth. To investigate whether *FLF* also has other roles in addition to its role in controlling the time of flowering, we examined its expression in a range of tissues. Figure 5B confirms the high expression of the *FLF* gene in vegetative rosette leaves, and reveals a strong expression in roots and lower expression in floral tissues, suggesting possible further roles for the *FLF* gene. No root phenotype has been observed in the transgenic lines. However, a number of lines had reduced fertility, which appeared to be caused by a lack of pollen in C24 lines or by abnormal carpels in Landsberg *erecta* lines. However, as the early Ac plants (see later) did not show these phenotypes it is unclear whether this is caused by change in expression of the *FLF* gene.

The expression of the *FLF* gene is markedly lower in post-vegetative tissues than it is in vegetative rosette leaves. We investigated the possibility that reduction in the level of expression of the *FLF* gene accompanies the transition to flowering. RNA was isolated from C24 and *flf* whole plants every 10 days post-sowing, until the stage where the majority of C24 plants had bolted (50 d). The expression of the *FLF* gene remained unaltered in these

plants (Figure 5C), suggesting that if reduction in the level of expression of the *FLF* gene does accompany the transition to flowering it must occur in only a few cells.

The *flf* mutant demonstrates a number of similarities to the *lhy* mutant described by Schaffer et al. (1998): they are both (semi-)dominant late-flowering mutants caused by insertion of foreign DNA adjacent to the gene. In wild-type the *LHY* gene is expressed for only a few hours around dawn, whereas in the *lhy* mutant, the *LHY* gene is expressed around the clock. Because of the similarities between *flf* and *lhy*, we examined the expression of the *FLF* gene in C24 and *flf* tissue harvested at dawn and dusk of an 8 h photoperiod. Although there was some difference in the expression of the genes between the two timepoints, there was no alteration of this pattern in the mutant.

Example 15 The Late-Flowering Phenotype Of The *flf* Mutant Is Suppressed By Vernalization Or By Gibberellic Acid Treatment

In a number of late-flowering mutants and ecotypes, low temperature treatment of germinating seed (vernalization) induces early flowering (Napp-Zinn, 1985), with a 4°C treatment for 21 days saturating the vernalization requirement to produce the shortest time to flower (Bagnall, 1992). The effect of vernalization on the time to flowering of hemizygous and homozygous *flf* mutants is shown in Table 2.

Table 2

Flowering Time of Vernalised C24 and Mutant Plants

| | LENGTH OF VERNALISATION | | |
|---|---|---|---|
| | 0 weeks | 4 weeks | 8 weeks |
| C24 (wild-type | 25.2 ± 0.2 | 13.6 ± 0.3 | -- |
| Hemizygous *flf* | 71.4 ± 1.2 | 39.3 ± 3.7 | 20.2 ± 0.9 |
| Homozygous *flf* | >150 | 100.8 ± 10.7 | 17.6 ± 1.3 |

Twenty seeds of the *flf* mutant and wild-type C24 were grown aseptically on MS medium in test-tubes, and exposed to either 4 or 8 weeks at 4°C. Non-vernalized plants were grown in soil (20 plants per 20cm pot). All plants were then grown at 23°C under fluorescent lights (16 h light, 8 h dark). The data are presented as the average number of days (± standard errors) until stem elongation, excluding the period of vernalization.

A 28 day treatment at 4°C resulted in a substantial reduction in the flowering time. However, eight weeks at 4°C was required to saturate the vernalization response in both hemizygotes and homozygotes to give a flowering time similar to that of the C24 control. This implies that there is an interaction between *FLF* gene expression and a component of the vernalization-induced pathway.

RNA was extracted from 12 d old C24 seedlings that were either vernalized or were untreated controls, and probed with *FLF* gene-specific probe. Figure 5E shows a dramatic decrease in *FLF* expression in vernalized seedlings compared with unvernalized seedlings, suggesting that a component of the vernalization signalling pathway controls *FLF* gene expression. Day 1 is the day on which seeds were transferred to the growth room. In *flf* mutant plants the level of transcript was reduced in 3 week vernalized seedlings, but not to the low levels observed in C24,

consistent with its only partial earlier-flowering character. *flf* mutant plants are being vernalized for 8 weeks to see whether they show an even further reduction in *FLF* gene expression.

As with other late-flowering vernalization-responsive mutants and ecotypes of *Arabidopsis*, *flf* mutant plants responded to applications of gibberellic acid ($GA_3$) by flowering earlier. Four week-old *flf* homozygotes treated with 1 μg of $GA_3$ every second day for a total of two weeks flowered two weeks after the final $GA_3$ application, compared to later than 20 weeks for untreated plants. However, a single treatment of 4 week old plants with 1 μg of GA3 was not sufficient to induce flowering of *flf*, although this amount of $GA_3$ induced early flowering of the late-flowering *fca* mutant (Bagnall, 1992), suggesting that *flf* requires a greater amount of $GA_3$ for floral induction. In contrast to the dramatic effect of vernalization on the expression of the *FLF* gene, exposure of either C24 or *flf* seedlings to $10^{-5}$ M $GA_3$ had no effect on the expression of the *FLF* gene.

Example 16 Movement Of An Ac Element Present Within The T-DNA Causes Alteration In Flowering Time

Two early flowering plants (M1 plants designated efSL3 and efSL4) were identified from one seedlot comprising bulked seed from *flf* mutant plants. Both plants flowered after 18 days, earlier than C24 which flowered after 30 days. PCR analysis using primers from within the T-DNA sequence and flanking genomic sequence confirmed that these early-flowering plants were derived from the *flf* mutant, and were not contaminants.

The tandem T-DNAs present in the *flf* mutant each contain an Ac element, and we considered the possibility that movement of Ac was the cause of the early-flowering phenotype. DNA was isolated from individual early-flowering M2 progeny of efSL3 and efSL4, digested with EcoR1 and probed with the 3' region of Ac. Figure 6A shows

the appearance of a new 2.1 kb band in the early-flowering plants, indicating the movement of Ac. The maintenance of the two original Ac bands present within the T-DNAs indicates that the Ac elements have remained in their original positions as well.

In order to determine the new location of Ac, we probed a similar DNA gel blot with probe 4 (Figure 2C), revealing a change in size of the 2.7 kb EcoRI fragment containing the promoter, MADS domain and part of intron I of the *FLF* gene (Figure 6B). The size of the 3' Ac fragment (2.1 kb) indicated that the Ac element had inserted near an EcoRI site, with the 3' end of the Ac nearest the EcoRI site. PCR using a primer in the 3' end of the Ac and primers near the EcoRI sites at either end of the 2.7 kb EcoRI fragment revealed that Ac had inserted within intron I in the *FLF* gene of efSL3. Sequencing of the PCR product determined the precise insertion point of the Ac element (Figure 6C). RNA was isolated from rosette leaves of early-flowering M2 plants and comparable-sized rosette leaves of C24 and *flf* plants, and probed with an *FLF* gene-specific riboprobe. As shown in Figure 5d, *FLF* gene expression was reduced to approximately 5% of the C24 expression level. It appears that the presence of Ac reduces the splicing efficiency of the transcript, hence reducing the amount of the normally-spliced product, and resulting in early-flowering.

Example 17 Excision of Ac From Intron I Causes Later Flowering and Increased *FLF* Transcript Levels

Twenty progeny of the M1 efSL3 plant were grown and their flowering time recorded. 15 M2 plants flowered at 18 days, the same as their M1 parent; however, 5 plants flowered later than their parent. These later-flowering plants were termed "neo-lates" to distinguish them from the original late-flowering *flf* mutant plants. Their approximate days to bolting were: 50 d (3.2), 50 d (3.4), 60 d (3.5), 85 d (3.5), 100 d (3.1). In the M2 progeny of

efSL4, 36 plants flowered at the same time as their M1 parent but one neo-late plant, 4.6, flowered at 38 d. DNA was isolated from individual plants and probed with 3′ Ac (Figure 7A). All the neo-lates appeared to be hemizygous for the presence of the 2.1 kb band, indicating hemizygosity for the presence of Ac in intron I. In some cases (3.1, 3.2, 3.3) the Ac element had relocated to a new site, as indicated by the appearance of the new band in Figure 7A.

To demonstrate that in each neo-late plant one copy of the Ac in intron I had excised, PCR was performed using primers derived from intron I sequence, flanking the site of Ac insertion. In each case a PCR product was generated, indicating that at least one copy of Ac had excised. The presence of the 2.1 kb band in Figure 7A indicates that at least one copy of Ac remains in intron I; thus each plant is hemizygous for the presence of Ac in intron I

The neo-late plant 4.6 is homozygous for the loss of the Ac element within the T-DNA closest to *FLF* (gene B). We believe that the sector of the M1 early-flowering parent from which the 4.6 seed derived must have been hemizygous for the presence of this Ac element, and plant 4.6 is a homozygous segregant for the loss of this Ac element. Presumably the loss of this Ac element was an event independent of that which resulted in the insertion of Ac in intron 1, as M2 progeny of the M1 plant efSL3 are homozygous for the presence of Ac at this location.

Total RNA was extracted from approximately 80 d old rosette leaves of the neo-lates and the gene B expression level was compared to that of 80 d rosette and cauline leaves of *flf* and 25 d rosette leaves of C24. In all neo-lates the transcript level was higher than that of the early-flowering parent, although lower than that of the *flf* mutant, consistent with them being hemizygous for the late allele.

Thus insertion of an Ac element into intron I of *FLF* (gene B) greatly reduces *FLF* (gene B) transcript levels and causes early flowering, and excision of Ac from intron I restores expression of the *FLF* (gene B) transcript and results in later flowering. This provides compelling evidence that over-expression of *FLF* (gene B) in the late-flowering *flf* mutant is the cause of the late-flowering phenotype.

<u>Example 18</u> <u>The Expression of the *FLF* Gene is Controlled by Known Flowering Time Genes</u>

In order to further understand the role of the *FLF* gene in the control of flowering we examined its expression in a range of *Arabidopsis* ecotypes and late-flowering mutants. Figure 8A shows the expression of the *FLF* gene in a variety of ecotypes. Interestingly the *FLF* gene is highly expressed only in ecotypes which have a late allele at the *FRI* locus (Pitztal and C24), and is particularly highly expressed in C24, but not in ecotypes with an early allele at the FRI locus (Columbia, Ws, Landsberg *erecta*).

To investigate this further we looked at the expression of the *FLF* gene in the Landsberg *erecta* ecotype with late alleles of either *FRI* or *FLC* (Figure 8B). Again *FLF* is expressed in the line of Landsberg *erecta* with the late $FRI^{Sf2}$ allele; however, it is also expressed in lines which contain late $FLC^{Sf2}$ and $FLC^{Col}$ alleles. The expression of the *FLF* gene in Landsberg *erecta*-$FLC^{Col}$ is interesting, as this plant has the same *FRI* and *FLC* genotype ($FRI^{early}$, $FLC^{late}$) as the ecotype Columbia, yet there is expression in Landsberg *erecta*-$FLC^{Col}$, but not in Columbia. This suggests that Landsberg *erecta* and Columbia differ in a third, unknown, locus, and that this locus in conjunction with late alleles at the *FLC* locus is able to induce expression of the *FLF* gene, in the absence of a late allele of *FRI*.

Many of the late-flowering mutants are in the Landsberg *erecta* ecotype, and we looked to see whether the

*FLF* gene is upregulated in these mutants. Figure 8C shows that the *FLF* gene is upregulated in the *fca* and *fve* mutants and slightly upregulated in the *fpa* and *fd* mutants in the Landsberg *erecta* ecotype and in *ld* in Ws ecotype, but not in any of the other late-flowering mutants tested.

We also looked at the expression of the *FLF* gene in the mutants *vrn1* and *vrn2*, which have reduced response to vernalization. Both mutants were isolated in the *fca* mutant background, and as *fca* and *vrn2* are closely linked only the *fcavrn2* double mutant is currently available. No increase in expression of the *FLF* gene in *vrn1* was detectable, however, the *vrn2fca* double mutant had increased in *FLF* gene expression over the *fca* mutant level. These data demonstrate that the function of the wild-type alleles of the *FCA*, *FVE*, *FPA*, *LD*, *FD* and *VRN2* genes is to down-regulate the *FLF* gene.

The expression of the *FLF* gene in early-flowering plants with reduced levels of methylation (Finnegan *et al*, 1996), is reduced suggesting that methylation may play a role in controlling the expression of the *FLF* gene, or a gene which is a regulator of *FLF*.

Vernalisation or some component in the vernalisation signal transduction pathway acts either to suppress *FLF* transcription or to increase *FLF* mRNA degradation. C24 has a strong vernalisation response, with plants vernalised for 3 weeks flowering in about half the time of unvernalised plants. A four week vernalisation period decreases *flf* flowering time somewhat, but an eight week period is required to bring flowering back to C24 times. This incomplete effect of short vernalisation periods on flowering time of the *flf* mutant correlates with the incomplete decrease in transcript levels in *flf*. This suggests that the higher level of the *FLF* transcript, and presumably of the *FLF* protein, titrates out the promoter of flowering produced in response to vernalisation. Longer

periods of vernalisation may produce more of the promoter, which might overcome the higher *FLF* transcript level.

Other ecotypes such as Ws, Landsberg *erecta* and Columbia show little response to vernalisation, under similar conditions to those which give good response with C24, the other ecotypes flower only 1 or 2 days early. We note that these ecotypes have very low *FLF* transcript levels, which cannot be decreased much further by vernalisation.

Pitztal is a late-flowering ecotype which has a strong vernalisation response, and has high *FLF* transcript levels. Vernalisation is expected to decrease the level of *FLF* transcript in these plants.

The *flf* mutant requires prolonged GA treatment to cause it to flower early. This suggests that the high levels of *FLF* transcript (and presumably FLF protein) may act to remove GA. Since *FLF* is a MADS-box transcription factor, it may do so by activating genes involved in the catabolism of GA, either directly or indirectly.

The lack of effect of GA treatment on the level of *FLF* transcript suggests that GA acts either downstream of *FLF*, or via another pathway. *ie*.

VERNALISATION → → decrease in *FLF* transcript → → GA → → FLOWERING

or

VERNALISATION → → decrease in *FLF* transcript → → FLOWERING

?→ → GA → → FLOWERING

with *FLF* normally acting to block the pathway between VERN and GA

*Tissue-Specific Expression*

Levels of expression of *FLF* in different tissues of the C24 ecotype were examined. High expression was observed in vegetative leaves, whereas expression in root was approximately ¼ that of leaves. There was a low level of expression in all tissues of the floral bolt, *ie*.

cauline leaves, bolt stem, flower buds, mature flowers, and a very low expression level in green siliques. RNA was isolated from rosette leaves and "apex", *ie.* the tissue remaining after as many as leaves as possible and the roots were removed; this tissue includes very small leaves and the apical meristem. There was no difference in expression level between these two tissue types.

In the *flf* mutant expression of *FLF* was twice that of C24 in vegetative tissue, and expression in floral tissues was relatively greater (about 3 times C24 level).

In the Columbia ecotype the level expression vegetative leaves was very low compared to C24, while in floral tissue there was about the same level of expression as C24.

This suggests that there may be a separate control of vegetative and floral transcription of the *FLF* gene.

*Developmental Expression*

As *FLF* appears to be a repressor of flowering, one prediction about its pattern of expression is that it might decrease prior to, or accompanying, the transition to flowering. RNA was extracted from whole plants that were grown on MS medium and harvested every 10 days. Under these conditions 50% of C24 plants were bolting after 50 d. There was no change in the level of *FLF* transcript in either C24 or the *flf* mutant. This suggests that if there is a decrease in the level of *FLF* transcript accompanying the transition to flowering it must occur in very few cells. This result also suggest that there is no decrease in the level of transcript in older leaves, *ie.* the transcript is not diluted out as the leaf grows.

*Circadian Response*

C24 and *flf* plants were grown in 8 h fluorescent photoperiod for 21 d and either maintained in this condition, or transferred to either continuous light or

continuous dark. RNA was extracted from plants harvested at either the start or the end of what would have been the 8 h photoperiod. In each case there was slightly higher expression at the earlier time point, suggesting a subtle circadian response. There was no difference in the pattern of expression between the mutant and C24.

Example 19 Isolation of a *Brassica napus FLF* homologue

Low stringency screening of a *Brassica napus* genomic library with an *FLF* probe lacking the MADS box region resulted in the isolation of 18 strongly-hybridizing plaques out of a total of 72,000 screened. These plaques have been purified, and are ready to sequence. The low stringency conditions were: hybridization at 28°C in 50% formamide overnight and washed with a final wash of 0.1 x SSC, 0.1% SDS at room temperature.

It will be apparent to the person skilled in the art that while the invention has been described in some detail for the purposes of clarity and understanding, various modifications and alterations to the embodiments and methods described herein may be made without departing from the scope of the inventive concept disclosed in this specification.

References cited herein are listed on the following pages, and are incorporated herein by this reference.

REFERENCES

An *et al*
EMBO J., 1985 4 277-284

Armstrong, C.L., Peterson, W.L., Buchholz, W.G., Bowen, B.A. and Sulc, S.L.
Plant Cell Reports, 1990 9 335-339

Bagnall, D.J.
Aust. J. Plant Physiol., 1992 19 401-409.

Bechthold, N., Ellis, J. and Pelletier, G.
C.R. Acad. Sci. III, 1993 316 1194-1199.

Bevan, M.
Nucl. Acids Res., 1984 12 (22) 8711-8721.

Brumbaugh, J.A., Meddendorf, L.R. and Grone, J.L.
Proc. Natl. Acad. Sci. USA, 1988 85 5610-5614.

Chang, C., Bowman, J.L., De John, A.W., Lander, E.S., and Meyerowitz, E.M.
Proc. Natl. Acad. Sci. USA, 1988 85 6856-6860.

Chou, M.-L. and Yang, C.-H.
Plant J., 1998 15 231-242.

Christou, P., McCAbe, D.E. and Swain, W.F.
Plant Physiol., 1988 87 671-674

Crossway *et al*
Mol. Gen. Genet., 1986 202 179-185

Dean, C., Sjodin, C., Page, T., Jones, J. and Lister, C.
Plant J., 1993 2 69-81.

Devereux, J., Haeberli, P. and Smithies, O.
Nucl. Acids Res., 1984 12 387-395.

Dolferus, R., Jacobs, M., Peacock, W.J. and Dennis, E.S.
Plant Physiol., 1994 105 1075-1087.

Elledge, S.J., Mulligan, J.T., Ramer, S.W., Spottswood, M. and Davis, R.W.
Proc. Natl. Acad. Sci. USA, 1991 88 1731-1735.

Feinberg, A.P. and Vogelstein, B.
Anal. Biochem., 1983 13 6-13.

Finnegan, E.J., Lawrence, G.J., Dennis, E.S. and Ellis, J.G.
Plant Mol. Biol., 1993 22 (4) 625-633.

Finnegan, E.J., Peacock, W.J. and Dennis, E.S.
Proc. Natl. Acad. Scie., 1996 93 8449-8454

Fromm *et al*
Proc. Natl. Acad. Sci., 1985 82 5824-5828

Gleaves, A.P.
Plant Mol. Biol., 1992 20 1203-1207

Herrera-Estrella *et al*
Nature, 1983a 303 209-213

Herrera-Estrella *et al*
EMBO J., 1983b 2 987-995

Herrera-Estrella *et al*
*In:* Plant Genetic Engineering, Cambridge University Press, NY, 1985 63-93

Koncz, C. and Schell, J.
Mol. Gen. Genet., 1986 204 383-396

Koornneef, M., Hanhart, C.J. and Van der Veen, J.H.
Mol. Gen. Genet., 1991 229 57-66.

Koornneef, M., Blankestijn-de Vries, H., Hanhart, C., Soppe, W. and Peeters, T.
Plant J., 1994 6 (6) 911-919.

Krens, F.A., Molendijk, L., Wullems, G.J. and Schilperoort, R.A.
Nature, 1982 296 72-74

Langridge, J.
Aust. J. Biol. Sci., 1957 10 243-252.

Lawrence, G., Finnegan, E.J. and Ellis, J.
Plant J., 1993 4 659-669

Lazo, G.R., Stein, P.A. and Ludwig, R.A.
Bio/technology, 1991 9 963-967.

Lee, I., Michaels, S.D., Masshardt, A.S. and Amasino, R.M.
Plant J., 1994b 6 (6) 903-909.

Lister, C. and Dean, C.
Plant J., 1993 4 (4) 745-750.

Longemann, J., Schell, J. and Willmitzer, L.
Anal. Biochem., 1987 163 16-20.

McNellis, T. W., Mudgett, M. B., Li, K., Aoyama, T., Horvath, D., Chua, N.-H. and Staskawicz
Plant J., 1998 14 247-257.

Napp-Zinn, K.
In CRC Handbook of Flowering Vol. 1. A.H. Halevy ed. (Boca Raton, FL. CRC Press). 1985 pp 492-503.

Newman, T. *et al.*
Plant Physiol, 1994 106 1241-1255

Paszkowski *et al*
EMBO J., 1984 3 2717-2722

Pietrzak, M., Shillito, R.D., Hohn, T. and Potrykus, I.
Nucl. Acids Res., 1986 14 (14) 5857-5868

Sanford, J.C., Smith, F.D. and Russel, J.A.
Methods Enzymol., 1993 217 483-509

Sanger, F., Nicklin, S. and Coulson, A.R.
Proc. Natl. Acad. Sci. USA, 1977 72 5463-5467.

Schaffer, R.1, Ramsay, N., Samach, A., Corden, S., Putterill, J., Carré, I. and Coupland, G.
Cell, 1998 93 1219-1229

Shannon, S. and Meeks-Wagner, D.R.
Plant Cell, 1991 3 877-892.

Southern, E.M.
J. Mol. Biol., 1975 98 503-517.

Sung, Z.R., Belachew, A., Shunong, B. and Bertrandgarcia, R.
Science, 1992 258 1645-1647.

Valvelkens, D., Van Montagu, M. and Van Lijesbettens, M.
Proc. Natl. Acad. Sci. USA, 1988 87 5536-5540.

Weigel *et al*
Cell, 1992 69 843-859

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISATION
(B) STREET: LIMESTONE AVENUE
(C) CITY: CAMPBELL
(D) STATE: ACT
(E) COUNTRY: AUSTRALIA
(F) POSTAL CODE (ZIP): 2612

(ii) TITLE OF INVENTION: CONTROL OF FLOWERING

(iii) NUMBER OF SEQUENCES: 3

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7968 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Arabidopsis thaliana

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGATCCAAGA AATAATTTCA TATGGGGCAC AGTTAAAAAA AAAACAATAA AATGATAATA 60

GTAAGGCTTG AACTTGGGTT GATGTGAGGC ACTATTAAGT AAAAAGCCAT TGTACTACTT 120

ACATTTTAAC TACATAATGT TAACTTATAT AATATTTATT GAATTCAGTA TAGGGCACAT 180

GCCCTATCCA TGACTAACGT GAGTCCGCCC TGATAGCGAG TAAGAAACGA GCAAAGGAAT 240

GCAAATTATG TGAATACTAT ACACCAACAG TGTAGACATG TAGCTACAAT GCGGCAATGT 300

TAAAAATAAA CTCAAATTGG TTTGGAGGGA ACAACCTAAT GCTTATAAGT ACACTTTTGT 360

GGTAAATACA ATCCAATTGA AAGTCTTTGT AGGTTTGGTT TGGTCCAATG AAATTGTATG 420

CGAGTTGAGT AAAATAACCT TAGTTCAAAA CATTAGATAT GTAATGGTCT AGATACGATG 480

GTAGCCAAAG ATTTGGGTTA AATTTAGAAT AAAGATCAAA GAACAAATGA CTGACTTCCT 540

TATGTGTGTT TGTTTATGTA AAGTCTTAAC TCGTGTCTTG CCAAATTAAT AAAAAGGTGC 600

ATTATACAAT GAACTAGAGC TTGTTCTCAT CAAAATTTTT AACATTAAAA TAAGTTATTA 660

AAATTTTGTT GATTATATAT GATTTCAATA TTAAATTATA TTTTTTGCTT ATAGCATCAA 720

AACTTCTTGG CACAGCTCCG AGTGTTACTG AAATGTTTGT GTGGCTCCAA TAGAAAAGTT 780

AATACCAATC ATGAATCATC AATATCGTTA CAAAATCAGT AATCAGTTTC ACCCACTCCG 840

AGTGTTAGTG AAATGTTTGT GTGGCTCCAT TAAAAAAGTT AATACCAATC ACCAATATCG 900

TTACAAAATT TGTAATCAGT TTCACCCACT TCTTTTCGTC TTGTTACTCG ATATAGCCTC 960

ACCTTTCAAG TGTCTAATAT AACACTTTTT AACAAAATAT TACAAAACAA GAAAAATTAT 1020

AGTAATTAGA TTAATCATAT GTAATCTATT ACATATAGTA ATTTGTATTA GTATCGTTTA 1080

TTGTGTTACC ATTCAAACGG TATAATCTAT ATAATTATTA GCAAAACAAA ATATAGTTTT 1140

AGTAATTAAT AAAATTAATA TAATGATAGT GATATTCAGA ATGTGGTATC
TAATCATGTA 1200

AAAATATATA ATAATAAAAT TAGATAAAGA AGAAATTGGT AAAAAATAAT
TAATGAGATA 1260

TAAAAGGAAA ACAAGCTGAT ACAAGCATTT CAGCCAAAAA AAAACAAGCT
GATACAAGCA 1320

AAAAAGAATT AGTAACTTTG AGCTATTGCC ATATGTGTGG ACATTTAAGA
TTTCTTGTAT 1380

TTTAGAAATC TTATATATTT CCACAATATA TTTACTACTT TCCTTTATTA
TTTGTGTTAA 1440

TCTCCCGAAC ATTATTATTT CAATACTATC TGAAAAACAC ATTTTTTTTT
ATAAAATTTG 1500

ATGACGTAGG CGAGTGGTTC TTTGTTTTTA CTATGTAGGC ACGACTTTGG
TAACACCTAC 1560

TAATTAAGTG CCAAATTTTA AGTTTTGAGA AGTCGGAAGA GTTCAAACCA
GTTTTAGGTT 1620

TCGATATGTC TACATAGTTC AAAGATGATG TAGAGTGGAG GTTCTTTCTG
CAATAGTTCA 1680

ATCCGTATCG TAGGGGAGGA AAGATAGTTT TCATTTAGCA AAGAAAGTGA
AAACTAAGGC 1740

AATGCAAAAG TAGCAAAGAC GCTCGTCATG CGGTACACGT GGCAATCTTG
TCTTCAAAAC 1800

GCAACGTTTT TATTCACATA TTTGGTTTTT TTGCATCACT CTCGTTTACC
CCCAAAAAAA 1860

AAAAATATCT GGCCCGACGA AGAAAAAGTA GATAGGCACA AAAAATAGAA
AGAAATAAAG 1920

CGAGAAAAGG AAAAAAAAAA ATAGAAAGAG AAAACGCTTA GTATCTCCGG
CGACTTGAAC 1980

CCAAACCTGA GGATCAAATT AGGGCACAAA GCCCTCTCGG AGAGAAGCCA
TGGGAAGAAA 2040

AAAACTAGAA ATCAAGCGAA TTGAGAACAA AAGTAGCCGA CAAGTCACCT
TCTCCAAACG 2100

TCGCAACGGT CTCATCGAGA AAGCTCGTCA GCTTTCTGTT CTCTGTGACG
CATCCGTCGC 2160

TCTTCTCGTC GTCTCCGCCT CCGGCAAGCT CTACAGCTTC TCCTCCGGCG ATAAGTACGC 2220

CTTTTCCTTA CCTGGGTTTT CATTTATTCC CCCTTTTATC TTCTGTTTTG TGCTCTTTTA 2280

CTTTTCTGAG AAAATAAAAA TAAAAAAACA ATTAATATAC CGTTTGGTTT TTTTCCGGCG 2340

GATCTCTTGT TGTTTCTCGG TTCTGTGTTT GTTTGTGTTT TTTTCTGCGA CCATGATAGA 2400

TACATGAGAT AACCAAATTT AAGGAAGAAC AATGTCGTGA AGAAGCTTTT TAGCTTCTTA 2460

CTTTTGTTCA TTTCTCTCTC TATTTCTTAA AAAAAAAAAT TCTGCATGGA TTTCATTATT 2520

TCCTTGGAAA AAAATTGCAT GTCCTTCACG ATTTGTTTGA TACGATCTGA TGCGTGCTCG 2580

ATGTTGTTGA GTGAAGTTTC AAGCCATCTT TGATTGTTTC TTACCTTTAG AGATTCCTTA 2640

AGTTTTTGAA GAGTTAATTA TATATCACAA GACTAATGAT TAATGTCTCT TTTCAGAGTG 2700

ATTAAAATTC ATTGGATCTC TCGGATTTGT ATGCAATGCA CTTACGGGAG ATCTATAGAG 2760

TTGCTATGGG GTTAATGCTG AACAATGTAT ATATACCACA TTGTGCAGCT ATTGACTATA 2820

TAAGACTATT GTTAATCTTC TATGAATTCC TATCTTTGCT GTGGACCTAT TACTTGGTGA 2880

TTATCCAAAT TAGTGTTTTT AATTGATTCA TATTTTTCAT ACACAGTAGT TTTGAATTTT 2940

GGTAGCTTCA AAAAACTCAG CCTCACAATT AGTACTTACC GCACATATGC TACTTCAGTA 3000

ACATAATCTG GTTATCGATT GCGATTCTTT GAATCACAAT CGTCGTGTGC TATATATATA 3060

ACACCTTTTG CTGTACATAA ACTGGTCTAA TTTTAGACTA ATTAAATTTC ATTGTTCTCT 3120

TGGATTTGTA TATGCACGTC CGGGAGATTT ATAAATAAAA TTAGTATGAG GTTAATGGTA 3180

AAAAGGATCA AGAAGTTTGG TTTTAAATGT AAGCCACATT AATTGGGAAA CTATGACTAA 3240

AAGATGAATT GGTAATATAT ACATAATATT TTTAACGAAT TTCTCTCCTT TTTATGGGAT 3300

ATGCTATTTG AAGATTGTCC AAAGGTTTAT AGTTTCCCAC TCTTGCAGTT ACACACATAG 3360

ATTTGCCTCA TATTTATGTG ATTGTATATC AATTATCGCC CTTAATCTTA TCATCGTTGT 3420

GTTCAATTAT GACTTTGTTC CTATTCGTTA AAATTGACAA TCCACAACCT CAATCTTTTG 3480

TTGTGAAAAT CGACAATCAC ACAACCTTTG TATCTTGTGT CTTTTGTCAT GGAAATTGTC 3540

ATTCACACAG CCTTGTTTCT TTGGTGCCTC TAGGAAATTG AAAATCCCAC AACACTTGTC 3600

TTCATGTAAG AAATACCAAC CTCTTTGGTA CGGATCTATA ATGAATCAAT ATAATCCTAT 3660

ATATAAGTTG TCAAAATTGA ATCTGGTGTA GTGTCTACTA CAACCCTCCA ATATAATAAC 3720

CAAATGGTTG TAGTAGTTTG GCCATGTTGG TCAAGATCGC TGGCCGATTC TCACTTGATG 3780

CATACTTTGT TAGGATTTGT TCACCCCTAG TTAGGTCCAG CCTTGGAATT GTCGAGACAC 3840

CTGACTAGAA CTCCTGGTCT TAATTATGAT TTAATAAAGA AGAAGCCTTT TAGAACGTGG 3900

AACCCTTAGT TACTCAGTTA CTCTTTTTGC ATACTTAGGT TGATGCAAAG AGCTTAACTT 3960

CACAATAGGA CTGATATCTA TTAACAAAAC AAATTAAGTG AAGTTTTGTC AAAATTGTTG 4020

GATCTTCTAG GTCAATATGT AGTTTAGTTT TTATCTGTCT TAGTCGCTTC CTTCTATGGA 4080

AGATATATTT ATAGATATGT GTGATAAGTT TCCTACTAAT ATTAGTTAGT GTAACTTCAA 4140

GGGCAGAAAA CTCTTTACTT TTATTGCTTG ATTAATTTGG GGTTTAAATA TAGGAAATTG 4200

GAACCTCACA GTTTCTATAA ACGAGTATGT AATTGAATGT GAAATAACAA AATGGAAGAC 4260

CGGCTTCCTA TTCTTAGGAG TCTTTTGATA TTTGCAAAAA AAACATATAC AATAGAAATA 4320

TGAGTTTTGT CTAAGACTCG GTCCATGTAT TTGGAGTTTG GCTTCCTCAT ACTTATGGTT 4380

ATCTGGTTAC CGCCACATCA TCATTATCAT CTTATGGGTC ATCAATACTA GCTCTATCGC 4440

TGGAAAAAAC CTTGTCCTCA AGGTTCATTG AAAAATCCGA AAAGTTTTCT CGTATATGTT 4500

GATATGGTAT TACTTACAAA CAAAGAGCTG ATGTTACCAA TTTTGACACG AGATTACTAA 4560

TGAACTCATG AAAGAGGCGT TTTTAAAAAA TTCTTTTTAA AACTGGGATA CAAAAAGAAA 4620

AGAGGTAACT AATAATTTGA TACCATTGTT CGTAGTCCTG ATCAAATGTT ATAAGGGTAA 4680

ACATGATAGA AAATAGAGGG TAAATAGGTT TTGTTCTTAT AATGGTTTTG ATAACACGCT 4740

TTGTAAAGGA TATAGGTGTT TTTTGATGCT AAAAGTTGTG GTATGGATCA AAACCAAAAT 4800

GGAAGCTCTG AATCTCTGAT AGAGGTTGCA ATTAGAATTA TATAAGTTAA TTTGCAAATG 4860

AATTGGAAGC AGTCTTCCAC TATTTGCTAT TGTTAGGGAA GTCTTTCAGT TAATTTCAGA 4920

AAATTAAGAG AAATATGACT TTCTAGACTC AGTCTGTGTA CTTGGAATTT TACTTCGGTT 4980

TACTTCCATG TCATCACATT GTGGCTCATC AATATATGTG TGTATATACA TTCATGAGTA 5040

TATATGATTT CTGGAAAAAT AAAAATTGCT TGTTTGCATT TAAGATTGGG GCTGCGTTTA 5100

CATTTTATAT TGCATCAATT ATTTCAACAT AGATTCACAA ACATAAATGC ATAGAAACAA 5160

TCTGGACAGT AGAGGCTTAT GTTTAGGGTT CTTATGTACC TTAACTAGTT TGACTTTAAG 5220

TTAATCAAAG CCAGCGCTAT CACTAAACTT TATCTGTATG CCTTTGTATG ACTTTTCTTT 5280

GAGGGAAAAT GTCATTTTCA ATCTGCCGAA ATATATAATA AATACATGTT AGCCCACATA 5340

ATTCATTGGA TAACTAATCT TTGAGCAATT TTTGGTAAAT GTTTTGGTTC TTTTCTTTTC 5400

TTGAGAGAGA AAAAAAATAT CAGATATTAT TAAATATTGC TTACAAAGCT AAGAACAAGT 5460

TAAAACTTTT TTGAAAAAGT GGAAATTCAG ATGTGCTACT GCTTAAACAT GAATATTAAG 5520

ATTATTGTTT TTCTGAAATG TTACGAATAC TAGCGTGTTA TATATATGTA AAAGGTAAGG 5580

TGTTCTCTCA ATGTTTCATA GTTTCCAGTG GCCTTTTCAA GGGTTAGCTA GTAGTTTTGA 5640

TCCTAACATA TTTTTATTTT TTTTGTCATC TCTCCAGCCT GGTCAAGATC CTTGATCGAT 5700

ATGGGAAACA GCATGCTGAT GATCTTAAAG CCTTGGTAAT ACAAACATTT TGAATCTTTT 5760

CCCTGATGGA GTTTTATAAG GCGTAAATTT ACTATTAGTT TGCCGAGTGA TCCTAAATAT 5820

AAAATGAGGT GGTGGCTCCA CATGCATTAT GCATACCGCA ATTTTCATAG CCCTTGTCTT 5880

TTACCGCTTC TTCTGTCCCT TTTTCATGGG CAGGATCATC AGTCAAAAGC TCTGAACTAT 5940

GGTTCACACT ATGAGCTACT TGAACTTGTG GATAGGTTAG TACTACTAAC TAAGACTATA 6000

TTTGCTCTCC ACCTTTGATT ACAAAGGAAT TAGTTTTTTT TTTGTCAAAC TATGAATATA 6060

TGCAGCAAGC TTGTGGGATC AAATGTCAAA AATGTGAGTA TCGATGCTCT TGTTCAACTG 6120

GAGGAACACC TTGAGACTGC CCTCTCCGTG ACTAGAGCCA AGAAGGTAAG TTGATTTCGT 6180

AATGTCTACT CCTTTCTGAA TTTTGTTTGC TGAGAACAAC CGTGCTGCTT TTGTTTGTTG 6240

CAGACCGAAC TCATGTTGAA GCTTGTTGAG AATCTTAAAG AAAAGGTCAG ATATTTGCTA 6300

CCAATTTTAT TGTACATCAG ATATATCCTC TTCTGTGTTG TCTCTGTTAC TTTAAGTCTG 6360

CTTAACGAGC TTGCACACAT ATTTGCAACT TTCTTCATAT GTTTTGGATT CCAAATTCTG 6420

AAGTTGTTAG GTTTAGAAAC TTGATCGGTA ATTGCTGAAC ATTTTGATCT TTAAATCAGG 6480

AGAAAATGCT GAAAGAAGAG AACCAGGTTT TGGCTAGCCA GGTAACGAAA GCTACATTTC 6540

CTAAAAATAT ATATGCATAA CTAATAAGCA CTGCGTGTTG TGTGTCCAAT GTCCATGTAC 6600

ATGGACATAG ATACACACTC TTATGCTTGC AGATATATAT ATATATATAT ATAGTCAGTG 6660

CATTTCAATC ATTCACTAGT TAGCACTTTC CTGGTCTTGT ATAGTTGTAT TCTAGACAAT 6720

TCTTCTCAAG ATTAGGGCAT TTTGGTTGTT GGTAGTTTGG TTTATTAGGG TTAGTGAGAT 6780

TATTACTGAA TAAGAACAGA AATTTGATAA CGGCTGGTTA GAGTTAAGGG AAATCAGATG 6840

AAGTTATTTT TTTATTTTTT ATCGAGTATA AATTACATGA TTGCTATATC ATTTTACTAA 6900

ATTAAGAAAA AAAAATTCCG GTTGTTGGAC ATAACTAGGT TTTGGTTCTT CTTCTTCGTT 6960

TTTTTCATGT TAAAGTGTTT AATTAGGTTT TGGTTCATTT GGAGATTTAG GAACCTTTTA 7020

TAGTCTGGTT AAGTCTGGGT TTGGTAGAGA TTCAATAAGA TTTCTTGATT CTCTTCAGGT 7080

TATGGTCTGG TTCAGTCTAG TTTAGTTCAA TATTGGTTTC CTTGAAGGTT GTGTAAACGT 7140

TGTCTATATT TAAGTTAATC ACCTTTTAAC CAAAAAAAAA AGTTTATGGA CCGATTAGTT 7200

TTTTTTTTTT TTTTTTTTTG TGATGGTTAG GTTTGGATCC GAGTGGCTCA GTTCCAACTC 7260

CAAGTGTCTA GAAGTAGTGC TACTTTTACA TGCTATATAT AGGTTAGATT
ATAAATTATA 7320

AACTGGTAAA AGATTATAGA TACTGCTTCC AAACTTAAAA GCTTAAACAT
AAAGAACACA 7380

CAAATTATGA GAAAAATAAC CTTCTGTAGT GTTTTTTAAT GGTTGTTATT
TGGTGGTGTG 7440

AAAAAGATAT TCCTTGGATA GAAGACAAAA AGAGAAAGTG AATAGTGATT
TTGACCTATG 7500

ATTATCGTAC AGATGGAGAA TAATCATCAT GTGGGAGCAG AAGCTGAGAT
GGAGATGTCA 7560

CCTGCTGGAC AAATCTCCGA CAATCTTCCG GTGACTCTCC CACTACTTAA
TTAGCCACCT 7620

TAAATCGGCG GTTGAAATCA AAATCCAAAA CATATATAAT TATGAAGAAG
AAAAAAAAAT 7680

AAGATATGTA ATTATTCCGC TGATAAGGGC GAGGGTTTGT ACATCTTAAT
ACTCTCTCTT 7740

TGGGCAAGAG AGTTTGTGTG TGATACTTAA GTAGAGGGAA CTAAGTCAAT
ACTATCTGTT 7800

TTAAGAGAAA AGGTTGATGA ACTTTGTACC TTATTCGTGT GAGAATTGCA
TCGAGATCTT 7860

GAGTGTATGT GTTCTTCTCT TCTGTCAAAA ACTTGTGTTT GCTTCACAGT
GAAGAAGCCT 7920

ACGGCTTATT TTGCAACAGG GACGTGGCTC TCTCTCTCTC TCTGCGCG 7968

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 943 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Arabidopsis thaliana

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

CGAGAAAAGG AAAAAAAAAA ATAGAAAGAG AAAACGCTTA GTATCTCCGG
CGACTTGAAC 60

CCAAACCTGA GGATCAAATT AGGGCACAAA GCCCTCTCGG AGAGAAGCCA
TGGGAAGAAA 120

AAAACTAGAA ATCAAGCGAA TTGAGAACAA AAGTAGCCGA CAAGTCACCT
TCTCCAAACG 180

TCGCAACGGT CTCATCGAGA AAGCTCGTCA GCTTTCTGTT CTCTGTGACG
CATCCGTCGC 240

TCTTCTCGTC GTCTCCGCCT CCGGCAAGCT CTACAGCTTC TCCTCCGGCG
ATAACCTGGT 300

CAAGATCCTT GATCGATATG GGAAACAGCA TGCTGATGAT CTTAAAGCCT
TGGATCATCA 360

GTCAAAAGCT CTGAACTATG GTTCACACTA TGAGCTACTT GAACTTGTGG
ATAGCAAGCT 420

TGTGGGATCA AATGTCAAAA ATGTGAGTAT CGATGCTCTT GTTCAACTGG
AGGAACACCT 480

TGAGACTGCC CTCTCCGTGA CTAGAGCCAA GAAGACCGAA CTCATGTTGA
AGCTTGTTGA 540

GAATCTTAAA GAAAAGGAGA AAATGCTGAA AGAAGAGAAC CAGGTTTTGG
CTAGCCAGAT 600

GGAGAATAAT CATCATGTGG GAGCAGAAGC TGAGATGGAG ATGTCACCTG
CTGGACAAAT 660

CTCCGACAAT CTTCCGGTGA CTCTCCCACT ACTTAATTAG CCACCTTAAA
TCGGCGGTTG 720

AAATCAAAAT CCAAAACATA TATAATTATG AAGAAAAAAA AAATAAGATA
TGTAATTATT 780

CCGCTGATAA GGGCGAGCGT TTGTATATCT TAATACTCTC TCTTTGGCCA
AGAGACTTTG 840

TGTGTGATAC TTAAGTAGAC GGAACTAAGT CAATACTATC TGTTTTAAGA
CAAAAGGTTG 900

ATGAACTTTG TACCTTATTC GTGTGAGAAA AAAAAAAAAA AAA 943

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 196 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Arabidopsis thaliana

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Met Gly Arg Lys Lys Leu Glu Ile Lys Arg Ile Glu Asn Lys Ser
 1               5                  10                  15

Ser Arg Gln Val Thr Phe Ser Lys Arg Arg Asn Gly Leu Ile Glu
                20                  25                  30

Lys Ala Arg Gln Leu Ser Val Leu Cys Asp Ala Ser Val Ala Leu
                35                  40                  45

Leu Val Val Ser Ala Ser Gly Lys Leu Tyr Ser Phe Ser Ser Gly
                50                  55                  60

Asp Asn Leu Val Lys Ile Leu Asp Arg Tyr Gly Lys Gln His Ala
                65                  70                  75

Asp Asp Leu Lys Ala Leu Asp His Gln Ser Lys Ala Leu Asn Tyr
                80                  85                  90

Gly Ser His Tyr Glu Leu Leu Glu Leu Val Asp Ser Lys Leu Val
                95                 100                 105

Gly Ser Asn Val Lys Asn Val Ser Ile Asp Ala Leu Val Gln Leu
               110                 115                 120

Glu Glu His Leu Glu Thr Ala Leu Ser Val Thr Arg Ala Lys Lys
               125                 130                 135

Thr Glu Leu Met Leu Lys Leu Val Glu Asn Leu Lys Glu Lys Glu
               140                 145                 150

Lys Met Leu Lys Glu Glu Asn Gln Val Leu Ala Ser Gln Met Glu
               155                 160                 165
```

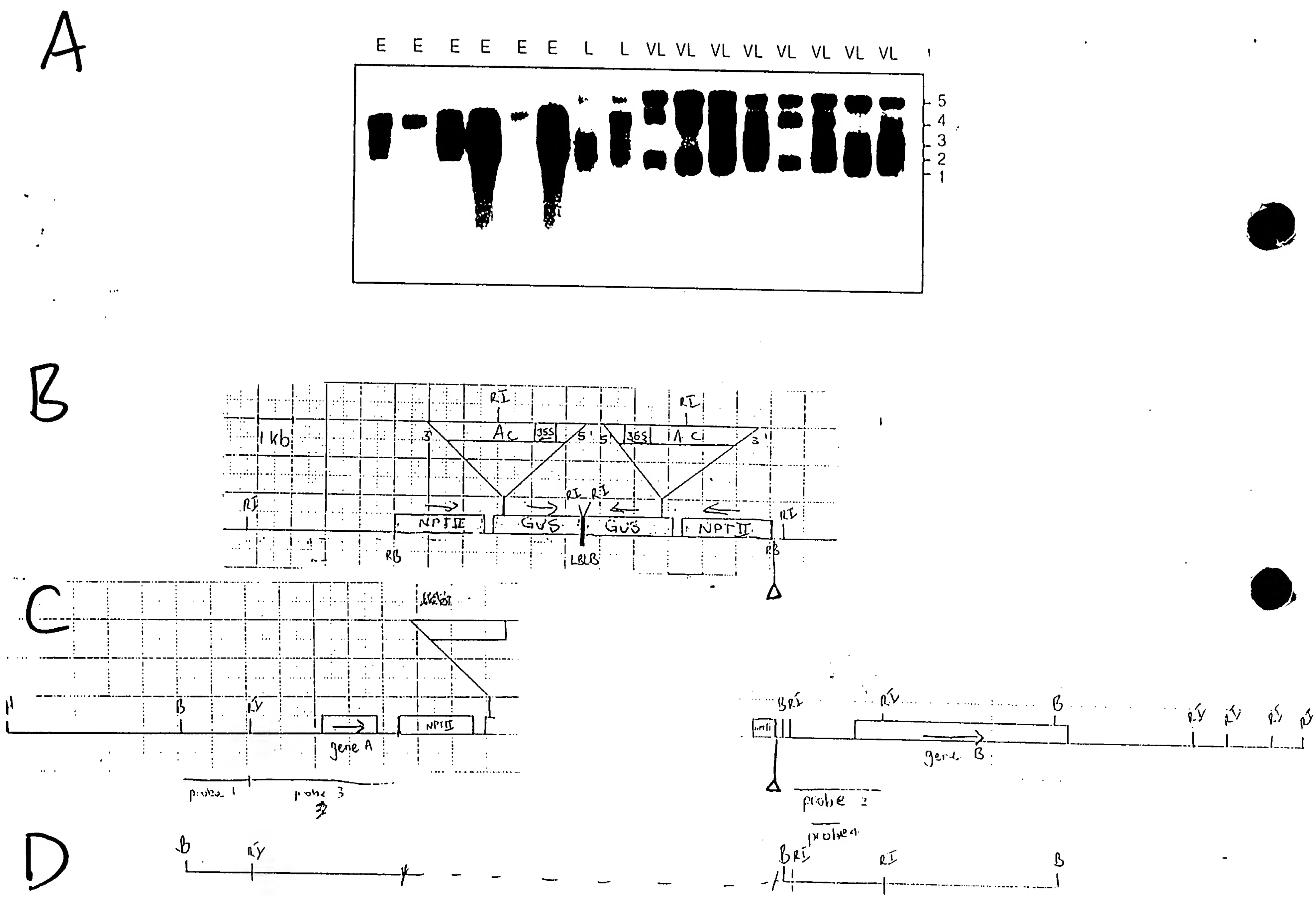
```
Asn Asn His His Val Gly Ala Glu Ala Glu Met Glu Met Ser Pro
                170                 175                 180

Ala Gly Gln Ile Ser Asp Asn Leu Pro Val Thr Leu Pro Leu Leu
                185                 190                 195

Asn
```

COMMONWEALTH SCIENTIFIC AND INDUSTRIAL RESEARCH ORGANISAITON

3 December 1998

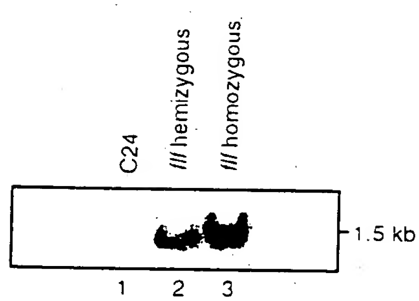
Figure 1

# FIGURE 2


A
E E E E E E L L VL VL VL VL VL VL VL VL
5
4
3
2
1
B
1 kb
RI
Ac
35S
AC
3'
5'
NPTII
GUS
RB
LBLB
C
gene A
probe 1
probe 3
B
RV
BRI
gene B
probe 2
D
RI

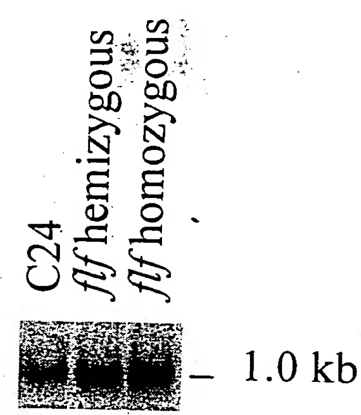
# FIGURE 3

A




Gene A transcript

B




Gene B transcript

Figure 4

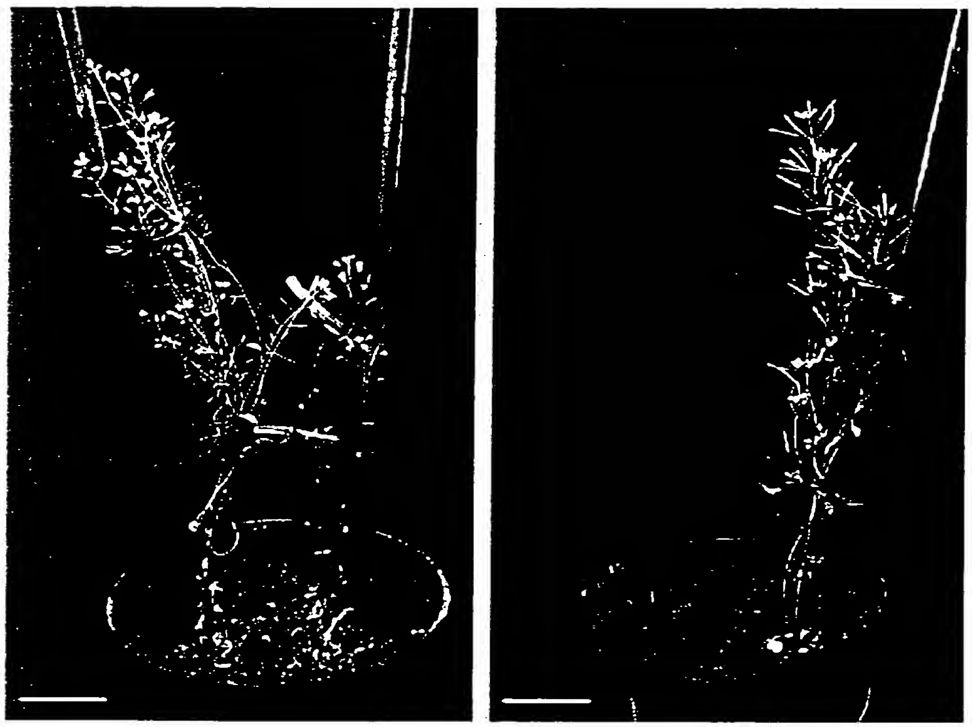
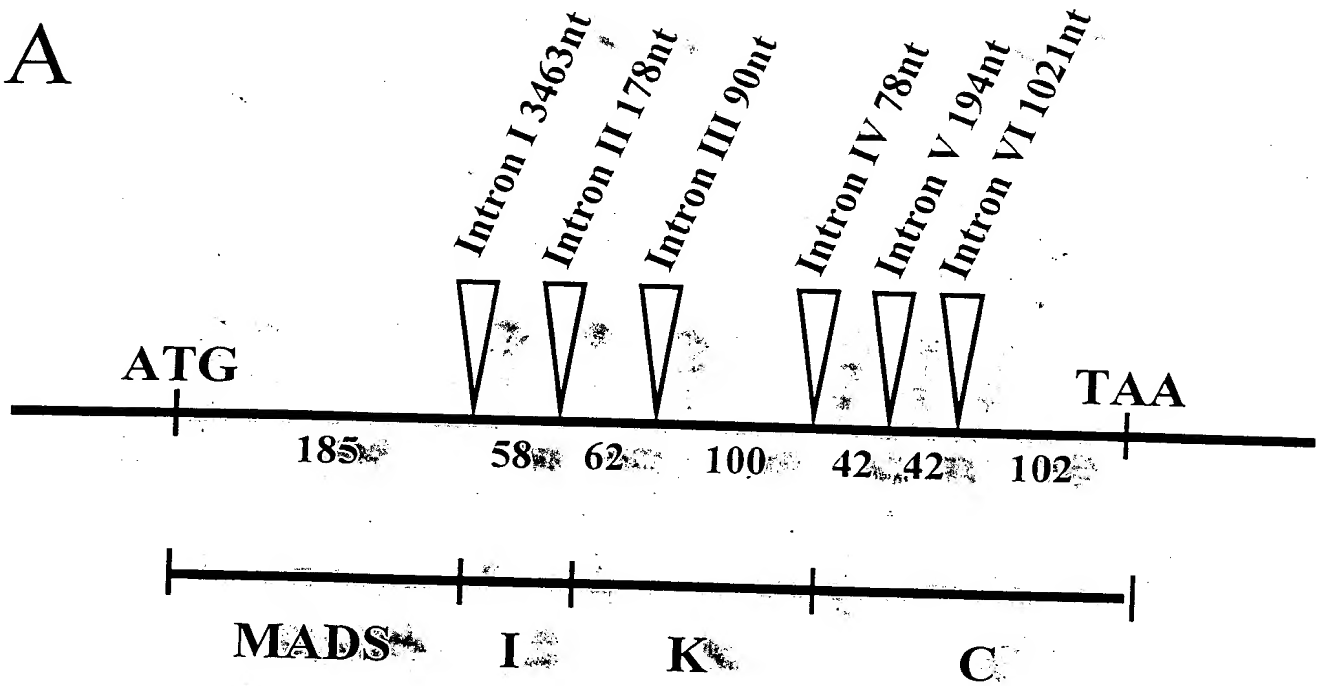


FIGURE 5(A)

A


Intron I 3463nt
Intron II 178nt
Intron III 90nt
Intron IV 78nt
Intron V 194nt
Intron VI 1021nt
ATG
TAA
185
58
62
100
42
42
102
MADS
I
K
C

# FIGURE 5(B-E)

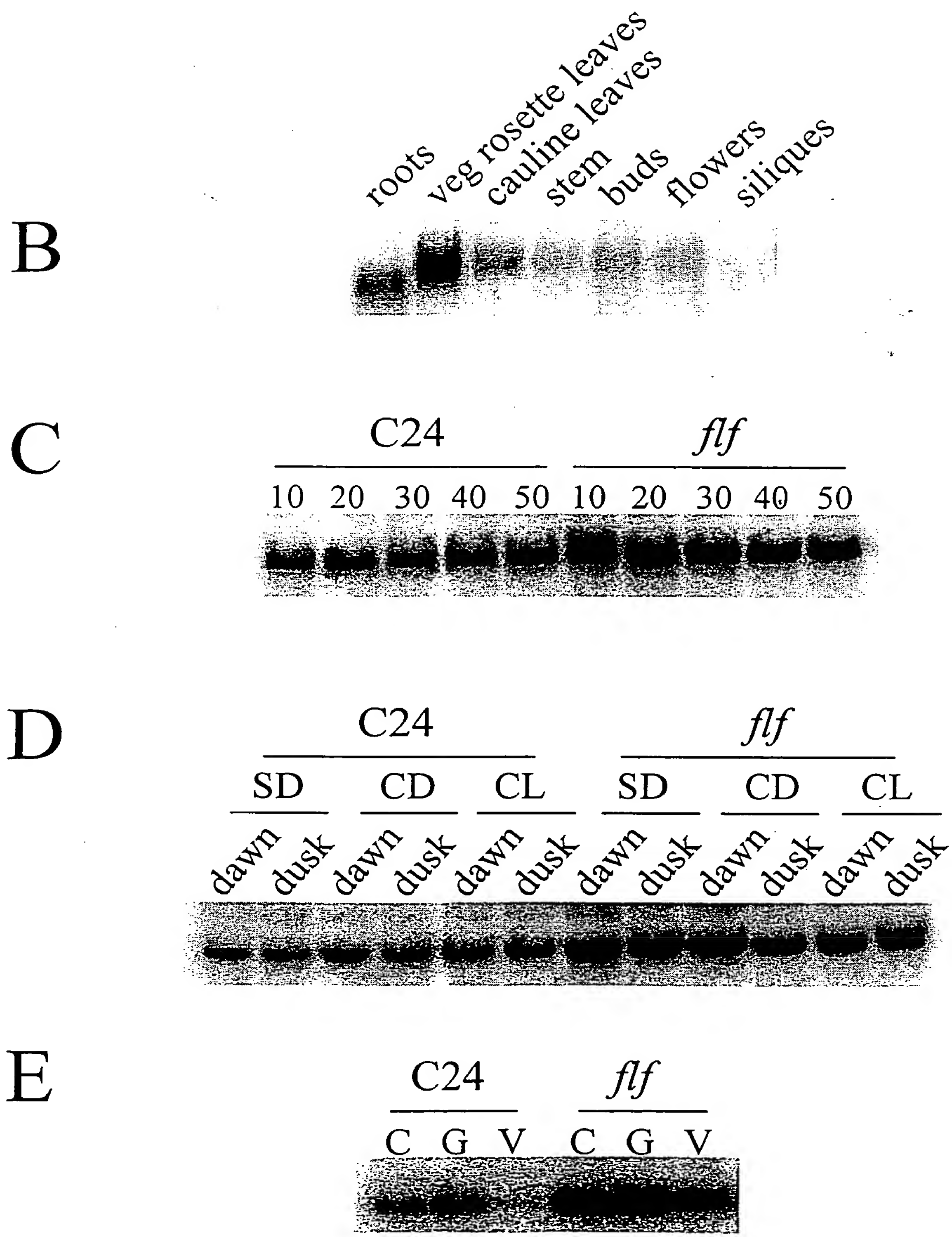
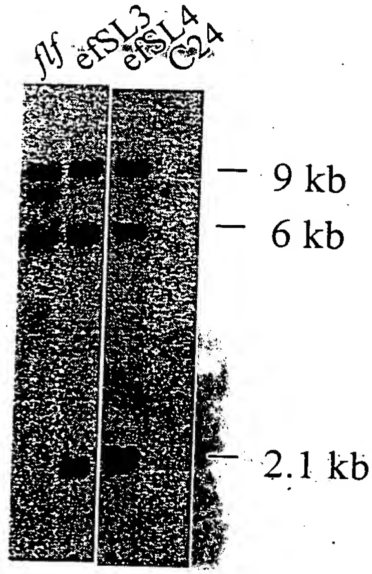

B
roots
veg rosette leaves
cauline leaves
stem
buds
flowers
siliques
C
C24
flf
10 20 30 40 50 10 20 30 40 50
D
C24
flf
SD CD CL SD CD CL
dawn dusk dawn dusk dawn dusk dawn dusk dawn dusk dawn dusk
E
C24
flf
C G V C G V

FIGURE 6(A,B)

A




B

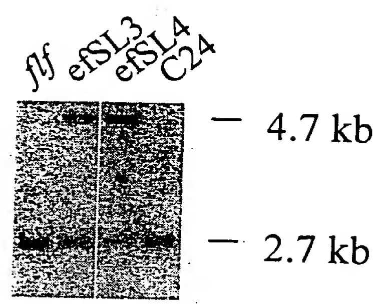
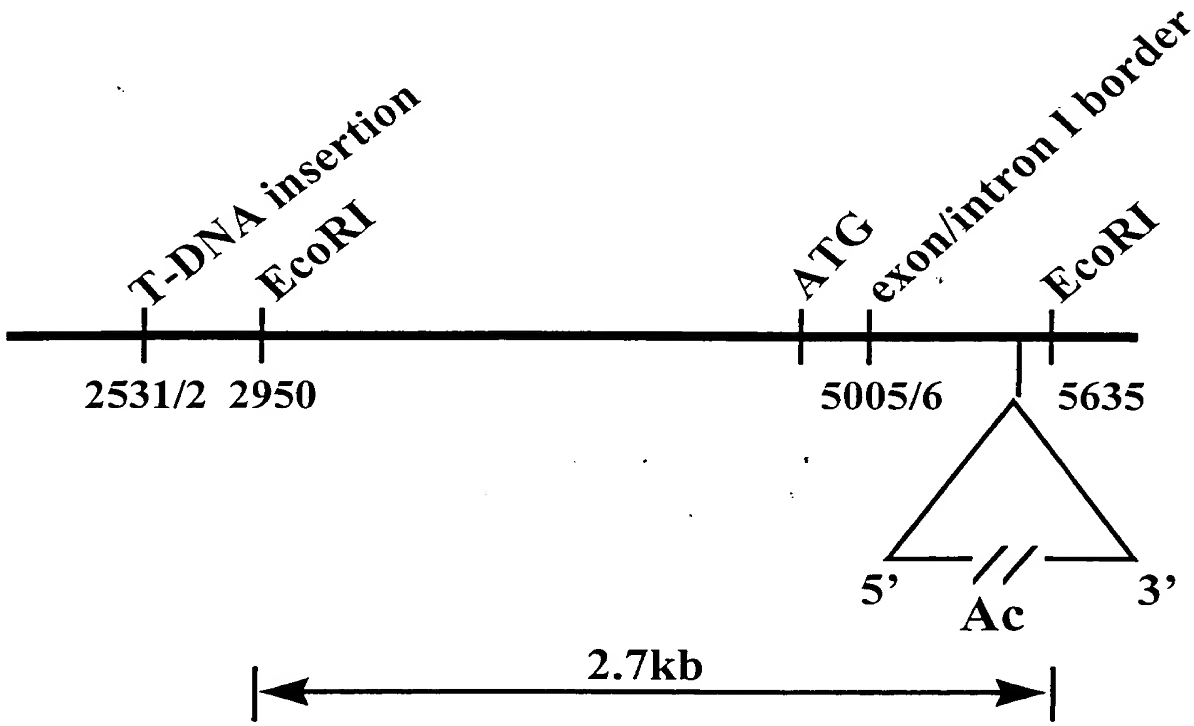
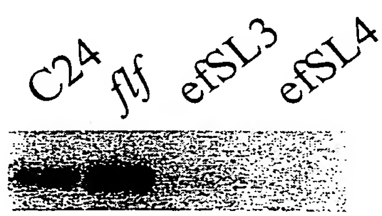


FIGURE 6(C,D)

C

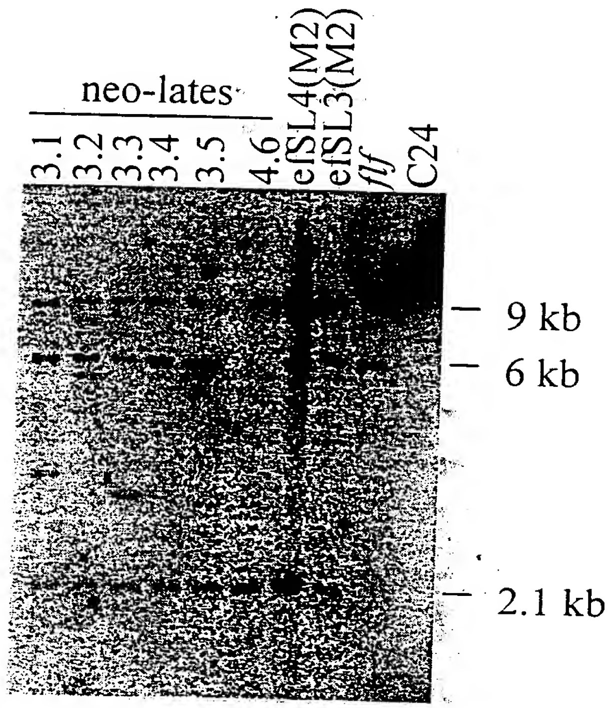



D



# FIGURE 7

A


neo-lates
3.1
3.2
3.3
3.4
3.5
4.6
efSL4(M2)
efSL3(M2)
*flf*
C24
9 kb
6 kb
2.1 kb

B

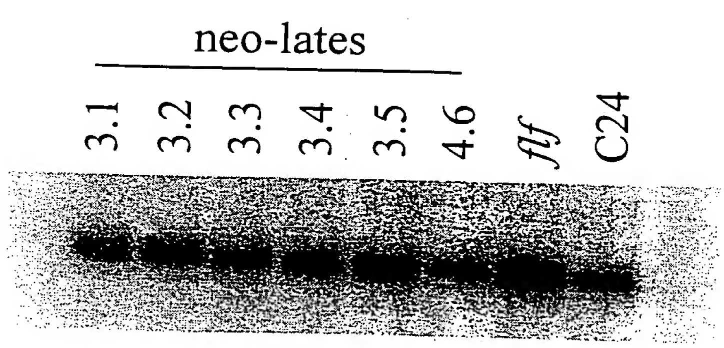
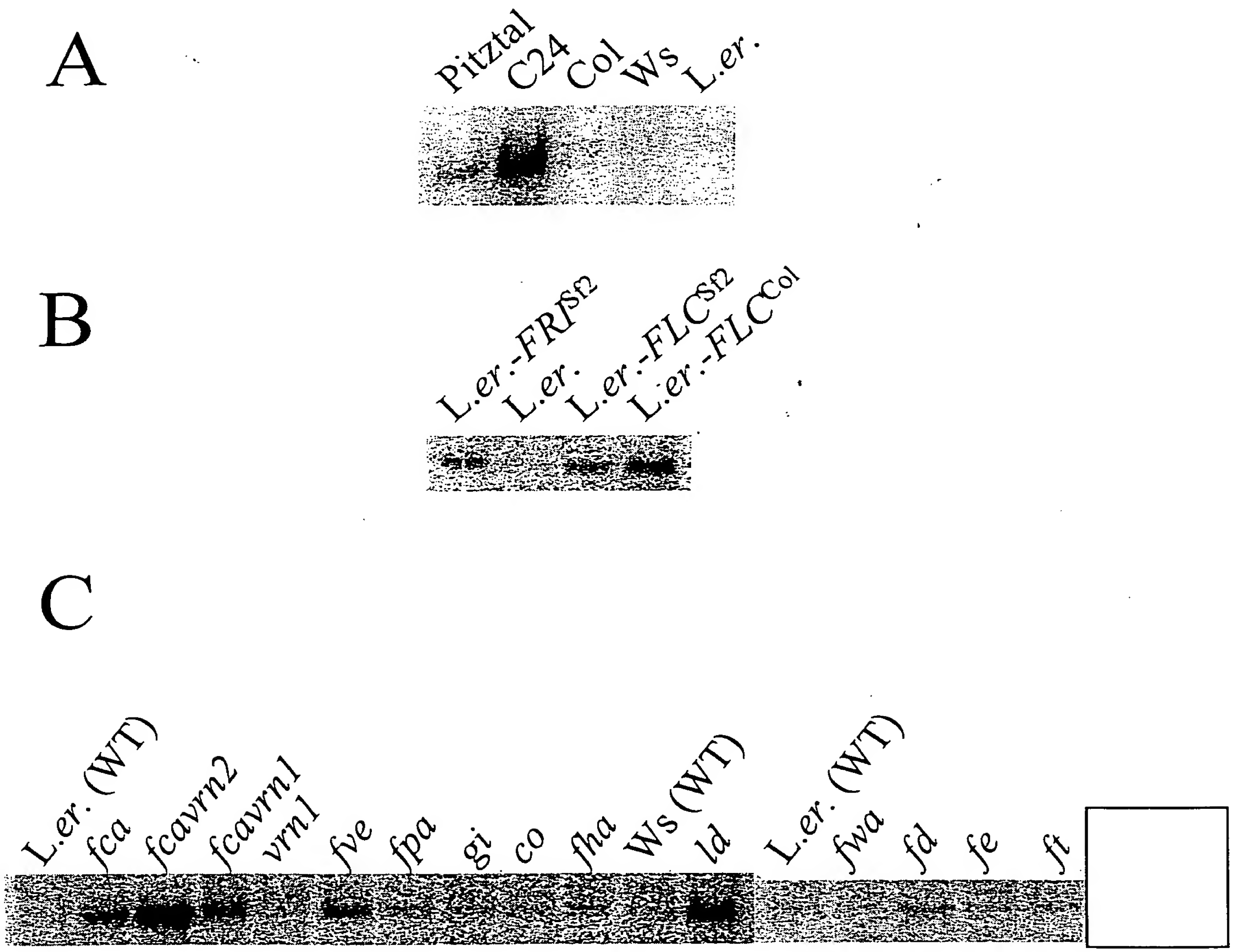

neo-lates
3.1
3.2
3.3
3.4
3.5
4.6
*flf*
C24

FIGURE 8



A
Pitztal
C24
Col
Ws
L.er.
B
L.er.-FRI$^{Sf2}$
L.er.
L.er.-FLC$^{Sf2}$
L.er.-FLC$^{Col}$
C
L.er. (WT)
fca
fcavrn2
fcavrn1
vrn1
fve
fpa
gi
co
fha
Ws (WT)
ld
L.er. (WT)
fwa
fd
fe
ft